Conformationally-assisted chemical reactions using a split protein

Caroline Morris

Submitted in accordance with the requirements of

University College London, UCL

Degree of Doctor of Philosophy

MRC National Institute for Medical Research

Division of Physical Biochemistry

Mill Hill

London

NW7 1AA

I, Caroline Morris, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

_____

Acknowledgements

Firstly I must thank my supervisor, Dr John Offer, for giving me the opportunity to work in his lab and inspiring me to aim high. I would like to thank Dr Justin Molloy, Dr Willie Taylor and Dr Derek Macmillan for taking the time to examine my project and offer advice as my thesis committee. I would also like to thank Dr Fabienne Burlina for introducing me to ligation, Dr Steve Martin for his invaluable help with BLI and CD, Dr Geoff Kelly for the NMR spectroscopy, and Dr Claudia Solscheid for assistance with UV/VIS experiments.

Thank you to everyone in the group for all the help, advice and stimulating conversation throughout my PhD. Cheers to everyone in Physical Biochemistry for making the division such a fun place to work, and for providing lunch and coffee breaks full of laughter. Nice atmos altogether, really.

A huge thank you to all of my wonderful friends, both within the NIMR and out in the real world, I couldn't have done it without you guys. In particular I must thank Andy for the words of wisdom that come with age and for always letting me win at snooker, Gemma for being my study room rock and a source of reassurance for so many years, and Olivia for being the best friend anyone could ask for.

The biggest thanks of all go to my family for encouraging me to be the best that I can be. My parents have provided constant support and unwavering encouragement in all of my endeavours, I will be forever grateful for all that they have done for me. Thank you.

I dedicate this thesis to my grandparents.

Presentations of this work:

- Poster presentation at PPSG Early Stage Researcher Meeting 2012, Royal Society for Chemistry, London, UK: *Simplifying native chemical ligation with an N-acylsulfonamide linker*.
- Oral presentation at NIMR/UCL Student Symposium 2013, NIMR, London, UK: *Proximity-assisted protein labelling by acyl transfer*.
- Poster presentation at PPSG Early Stage Researcher Meeting 2013, Durham University, Durham, UK: *Native chemical ligation with CI2 using a peptide N-(Me)sulfonamide*.
- Poster presentation at the 33rd European Peptide Symposium 2014 for the Young Scientist Mini Symposium, Sofia, Bulgaria: *Two methods simplifying peptide ligation: use of an N-acylsulfonamide linker, and α-methylcysteine.*

The work detailed in chapter 3 has been published in:

Fabienne Burlina, Caroline Morris, Raymond Behrendt, Peter White and John Offer, *Simplifying native chemical ligation with an N-acylsulfonamide linker*, Chem. Commun., 2012, **48**, 2579–2581.

Additional publication:

Fabienne Burlina, George Papageorgiou, Caroline Morris, Peter D. White and John Offer, *In situ thioester formation for protein ligation using α–methylcysteine*, Chem. Sci., 2014, **5**, 766-770.

Abstract

The aim of the project was to develop a novel method of protein labelling based upon split proteins. The proposed method would label a target protein *via* a small peptide sequence expressed at its *N*-terminus that is selectively targeted by its partner peptide carrying a tag. Split proteins were considered effective for a proximity-assisted labelling method because of the specificity of the peptide fragments for one another, and the extensive background literature available.

The project has studied bovine ribonuclease A (RNase A), chymotrypsin inhibitor 2 (CI2) and calbindin D_{9k}, three well-documented split proteins. In each instance the sequence was divided into two peptides that were synthesized by Fmoc-SPPS, with the exception of the S-protein of RNase A which was commercially available. Conformationally-assisted ligation with a peptide N-(Me)sulfonamide was initially explored using literature precedent, and the affinity of the interaction at this site was also measured by bio-layer interferometry (BLI) or UV/VIS spectroscopy. With limited success in attempts with both RNase A and CI2, calbindin D_{9k} was chosen because of the high affinity reported in the literature between its peptides. Their interaction was assessed by BLI and their secondary structure in the presence of either calcium or a chelator was compared by circular dichroism. In this system conformationally-assisted ligation was observed along with proximity-assisted acyl transfer of a small peptide tag, with promising results. The peptides were also observed to enable click chemistry, when synthesized as a peptide with *C*-terminal azide and a peptide with *N*-terminal alkyne, with no requirement for a copper additive.

The results described are successful applications of conformationally-assisted chemical reactions, enabled by the specificity and high affinity imparted by the split protein peptides. The reactions are chemoselective and active at biologically relevant concentrations, and are encouraging for the future development of a novel protein labelling method.

Table of contents

Abbreviations

AAC	Azide-alkyne cycloaddition
AcN	Acetonitrile
Boc	t-Butoxycarbonyl
BLI	Bio-layer interferometry
BPTI	Bovine pancreatic trypsin inhibitor
CD	Circular dichroism
CI2	Chymotrypsin inhibitor 2
CNBr	Cyanogen bromide
Da	Dalton
DCM	Dichloromethane
DIC	1,3-Diisopropylcarbodiimide
DMF	*N,N*-Dimethylformamide
DMSO	Dimethyl sulfoxide
EDT	1,2-Ethanedithiol
EDTA	Ethylenediaminetetraacetic acid
eq	Equivalent
Fmoc	9-Fluorenylmethoxycarbonyl
GdnHCl	Guanidine hydrochloride
Hmb	*N*-2-Hydroxy-4-methoxybenzyl
HOBT	1-Hydroxybenzotriazole
HPLC	High performance liquid chromatography
LCMS	Liquid chromatography-mass spectrometry

MALDI-TOF MS	Matrix assisted laser desorption/ionisation time-of-flight mass spectrometry
Me	Methyl
MESNA	Sodium 2-mercaptoethanesulfonate
MPAA	4-Mercaptophenylacetic acid
NCL	Native chemical ligation
PBS	Phosphate buffered saline
RNase A	Bovine ribonuclease A
RP	Reverse-phase
RT	Room temperature
SPPS	Solid phase peptide synthesis
SPR	Surface plasmon resonance
TCEP-HCl	Tris(2-carboxyethyl)phosphine hydrochloride
TES	Triethylsilane
TFA	Trifluoroacetic acid
TMS-diazomethane	Trimethylsilyl-diazomethane
t = 0	Time zero
wt	Wild type

Amino acids

Alanine	Ala	A
Cysteine	Cys	C
Aspartic acid	Asp	D
Glutamic acid	Glu	E
Phenylalanine	Phe	F
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Lysine	Lys	K
Leucine	Leu	L
Methionine	Met	M
Asparagine	Asn	N
Proline	Pro	P
Glutamine	Gln	Q
Arginine	Arg	R
Serine	Ser	S
Threonine	Thr	T
Valine	Val	V
Tryptophan	Trp	W
Tyrosine	Tyr	Y

List of figures and schemes

List of tables

1. Introduction

1.1 Proximity-assisted protein labelling with split proteins

1.1.1 Overview

The aim of the project was to develop a novel method of protein labelling using the peptides of a split protein and a proximity-assisted chemical reaction. These peptides would provide specificity, with one peptide selectively targeting the other expressed on the protein of interest. Ideally the smallest peptide possible would be covalently attached to the target protein in order to minimally impact upon the target protein's activity. Split proteins were considered effective for a proximity-assisted labelling method because of the specificity of the peptide fragments for one another, their widespread use and the extensive background literature available. In addition to the flexibility in choice of small synthetic dyes or tags to that most suited for the technique used, such a labelling method would combine the selectivity of split proteins with the bioorthogonality of a chemical reaction.

1.1.2 Background

One of the most powerful and commonly-used methods of labelling involves expression of green fluorescent protein (GFP), a technique that has revolutionized the field of protein labelling (Tsien, 1998). However its size, 238 residues and ~ 27 kDa, has led to the search for alternative methods with smaller components to minimally interfere with the activity of the target protein (Thomas and Maule, 2000). There are many covalent tags that bind based on chemical reactions with residues within the target protein, Cys being particularly popular because of its high nucleophilicity. Cys residues occur rarely in nature, just 1.79 % in nuclear proteins (Fukuchi and Nishikawa, 2001),

and those present are typically buried or forming disulfide bonds rather than accessible at the surface. As such site-directed mutagenesis can be used to introduce a Cys to allow for site-specific labelling, though this Cys residue must be maintained in its reduced form and the effect of altering the sequence on the folding and activity of the protein verified. This method is better suited to *in vitro* labelling with purified proteins than to *in vivo* labelling due to a lack of specificity, though one method generates specificity by introducing a tetraCys motif into GFP for *in vivo* FLASH labelling (Tsien, 1998). The suggested split protein approach would overcome selectivity problems in a cellular environment by the affinity and specificity of the split protein peptide partners for one another. The proposed method would involve the expression of the acceptor peptide at the *N*-terminus of a protein. The termini of a protein are typically accessible, potentially due to their charged nature, their availability for post-translational modification, or due to thermodynamic and kinetic factors (Jacob and Unger, 2007). A number of labelling methods use the termini of the target protein, for example combining native chemical ligation at the *N*-terminus with oxime ligation at the *C*-terminus (Yi et al., 2011) and dual-labelling by expressed protein ligation (Cotton and Muir, 2000).

Labelling method	Category	Signature feature
HaloTag (Los et al., 2008)	Enzyme-based	Chloroalkane linker with fluorescent tag
Sortase (Tsukiji and Nagamune, 2009)	Enzyme-based	Recognises LPXTG motif
FLAG-tag (Nonaka et al., 2009)	Affinity-based	Recognised by Ni(II)-DpaTyr probe
CA6D4 tag (Nonaka et al., 2010)	Affinity-based	α-Chloroacetyl functionality of probe
Affibody (Holm et al., 2009)	Affinity-based	Acrylamide modification binds nucleophilic side chains
SpyTag-SpyCatcher (Zakeri et al., 2012)	Affinity-based	Spontaneous isopeptide bond formation
Coiled-coils (Wang et al., 2014)	Affinity-based	Thioether bond formation with α-chloroacetyl

Table 1.1: Overview of selected proximity-assisted protein labelling methods

Methods investigating proximity-assisted chemical reactions to avoid problems with selectivity were explored and are summarized in Table 1.1 and Figure 1.1. Enzymes proved to be popular, using labels such as in the HaloTag approach (Los et al., 2008). The HaloTag is a modified haloalkane dehalogenase that covalently binds to ligands consisting of a chloroalkane linker and a fluorescent tag, compatible with *in vivo* imagine (Kosaka et al., 2009). Though this area of protein labelling has increased specificity the tags involved can still be quite large. A popular method is that of sortase labelling, based on its recognition of the short peptide sequence LPXTG (Tsukiji and Nagamune, 2009, Antos et al., 2009, Yamamura et al., 2011). Methods such as this involve more discrete components, but still require the addition of enzymes.

In turn this work led to the development of affinity labelling techniques, spontaneous reactions that do not require the addition of enzymes. Proximity-directed acyl transfer for protein labelling explored the use of the FLAG-tag (DYKDDDDK) (Nonaka et al., 2009). Until then labelling a protein with the genetically-encoded FLAG-tag required the use of FLAG-tag antibodies, their cost and size often making their application problematic. This group pioneered the development of the Ni(II)-DpaTyr probe selective for part of the FLAG-tag sequence, modified to contain a thioester (Figure 1.1a). When the probe and tag interact and form a binding complex the thioester is brought into close proximity with a Lys side chain of the FLAG-tag, inducing acyl transfer to form a covalent bond. This group later went on to label G-protein coupled receptors (GPCR) containing the tag CAAAAAADDDD (CA6D4) with a Zn(II)-DpaTyr probe, the α-chloroacetyl modification of the probe reacting with the Cys of the tag (Nonaka et al., 2010). The FLAG-tag has previously been used to label GFP by way of protein trans-splicing with split inteins (Giriat and Muir, 2003). In an effort to overcome the limitation in some cases of antibody affinity an affibody, a non-immunoglobulin scaffold, was used to induce proximity-assisted covalent bond formation for cell surface labelling of the Z domain of *Staphylococcous aureus* Protein A (ZSPA) (Holm et al., 2009). Using a modified affibody with a weak electrophile, acrylamide, near the binding site allows for covalent bond formation with nucleophilic side chains in the target protein, only doing so upon interaction with the target protein (Figure 1.1b). An impressive example of affinity labelling is the SpyTag-SpyCatcher method where the split protein fibronectin-binding protein FbaB of *Streptococcus pyogenes* generates two components that rapidly form an amide bond (Zakeri et al., 2012). The second immunoglobulin-like collagen adhesion domain (CnaB2) of this protein is split at the CnaB2 triad (Lys31, Asp117 and Glu77) to produce a 13-mer

peptide termed SpyTag and its 113-mer protein partner SpyCatcher. These split protein fragments undergo rapid covalent bond formation, with a spontaneous isopeptide bond forming between Lys31 and Asp117 (Figure 1.1c). Coiled-coils have been shown to combine the stability of proteins with the cost-effective nature of peptide synthesis (Chao et al., 1998, Thomas et al., 2013, Sanchez-Ruiz et al., 2013, Ryadnov and Woolfson, 2007). In a recent example protein labelling is brought about by a proximity-assisted reaction between coiled-coil peptides, one with a Cys and the other with an α-chloroacetyl modification (Wang et al., 2014). The peptides were 21 residues in length, and the covalent bond formation upon assembly of coiled-coil pairs demonstrated *in vitro* labelling in addition to labelling of proteins at the cell surface. However one concern with the use of coiled-coils in a labelling approach is the potential for oligomerisation of the peptides.

A detailed review on the use of chemical tags in fluorescence imaging can be found by Wombacher and Cornish (Wombacher and Cornish, 2011).

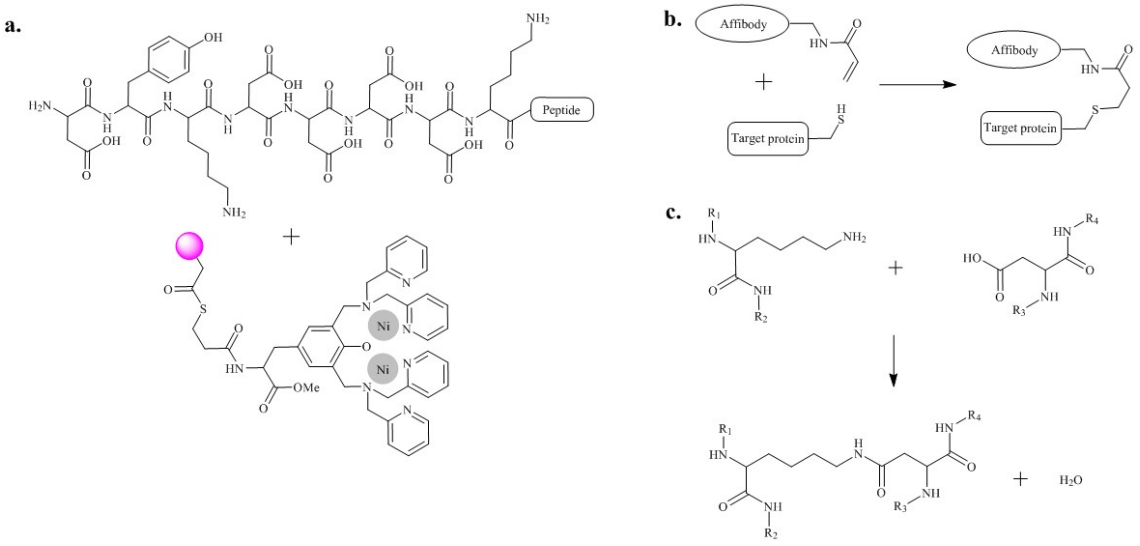


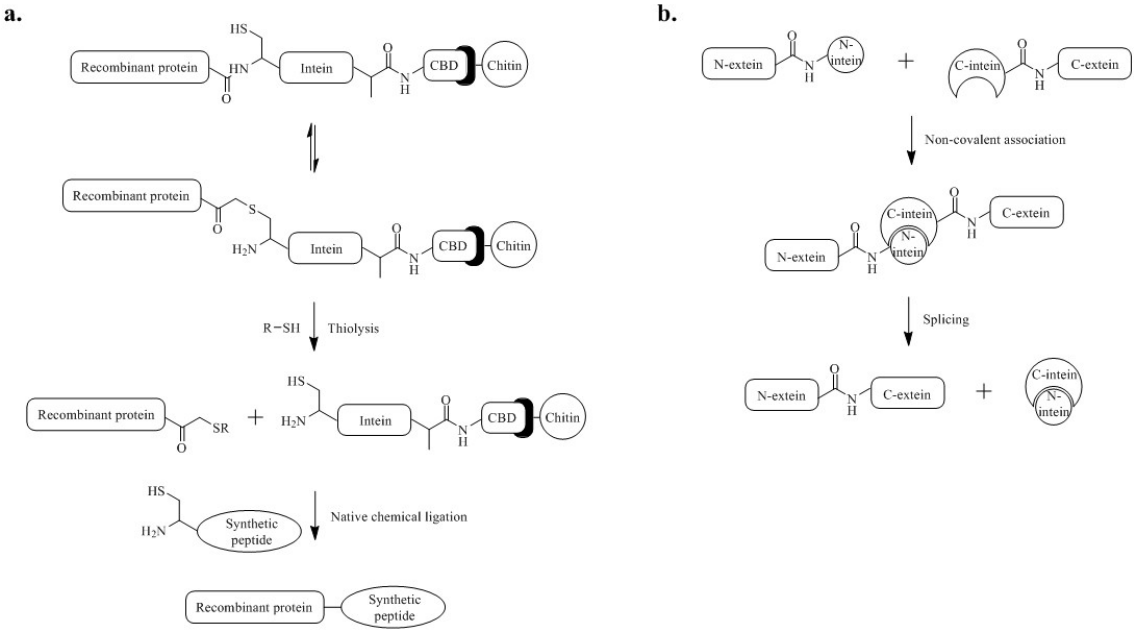
Figure 1.1: Overview of selected proximity-assisted protein labelling methods

a. FLAG-tag approach, with Ni(II)-DpaTyr probe. Adapted from Nonaka et al, 2009.
b. Affibody approach, with nucleophilic side chain recognised by acrylamide. Adapted from Holm et al, 2009.
c. Isopeptide bond formation in the SpyTag-SpyCatcher approach. Adapted from Zakeri et al, 2012.

1.2 Split proteins

1.2.1 Introduction to split proteins

Split proteins are those that are separable into peptides that can reconstitute *via* a non-covalent interaction with a high and specific affinity, reconstitution being well-studied in a number of systems as reviewed by Linse and colleagues (Carey et al., 2007). Split proteins are commonly used in the fields of cell and chemical biology, particularly the use of split inteins (Mootz, 2009, Shah and Muir, 2014). Split inteins have found uses in techniques for peptide cyclization (Scott et al., 1999), along with the preparation of semisynthetic proteins by protein trans-splicing (Muralidharan and Muir, 2006, Paulus, 2000, Giriat and Muir, 2003) and expressed protein ligation (Muir et al., 1998, Severinov and Muir, 1998, Muir, 2003) (Scheme 1.2). As with the proposed split protein approach, the affinity of the split intein fragments for one another allows for a specific and localised reaction (Shi and Muir, 2005). Inteins are internal protein domains that are able to excise themselves out of a precursor protein. This autocatalytic reaction was first identified from discrepancies between the size of mature yeast vacuolar ATPase and that predicted from its gene sequence (Hirata et al., 1990). In the development of split intein-based approaches, the mechanism by which naturally-split inteins associate has recently been elucidated (Shah and Muir, 2014). Muir and colleagues found that just the *N*-intein fragment is partially folded while the *C*-intein fragment is disordered, and that the disordered section is key to their interaction *via* a capture step that forms an intermediate. This subsequently collapses into the functional conformation, the native structure with the unique intein fold.



Scheme 1.2: Preparation of semisynthetic proteins by expressed protein ligation and protein trans-splicing

a. Expressed protein ligation involves a recombinant protein expressed *N*-terminal to an intein, complete with chitin binding domain (CBD) for purification, that is cleaved upon treatment with thiols. The recombinant protein thioester is then free to undergo native chemical ligation with a synthetic peptide containing an *N*-terminal Cys.
b. Protein trans-splicing uses a split-intein separated into *N*- and *C*-pieces that associate non-covalently to effect protein splicing, only doing so upon association with one another.

Adapted from Muir, 2003.

Split proteins play an important role in protein-fragment complementation assay, with the yeast two-hybrid system being a classic example. Yeast Gal4 is split into two domain proteins, both of which are essential for LacZ transcription. LacZ acts as a reporter gene, and the assayed proteins are fused to each of the Gal4 domains. If these proteins interact, Gal4 is reconstituted and thus LacZ is transcribed (Fields and Song, 1989). The split-ubiquitin two-hybrid system overcomes the solubility issues associated with the yeast two-hybrid system, and allows for the assay of interactions between insoluble membrane proteins (Stagljar et al., 1998). This is based on the development of split-ubiquitin for assaying *in vivo* protein interactions, where a reporter protein fused to the *C*-terminal ubiquitin fragment is cleaved upon reconstitution if the *N*-terminal ubiquitin fragment is present (Johnsson and Varshavsky, 1994). Additionally split-tobacco etch virus (TEV) monitors the return of proteolytic activity upon interaction of fused proteins (Wehr et al., 2006). Reassembly of split-GFP was first reported with GFP fragments non-covalently associated with leucine zippers (Ghosh et al., 2000), paving the way for the use of GFP reconstitution in protein-fragment complementation assay. Variants of split-GFP used, in a form of protein-fragment complementation assay known as bimolecular fluorescence complementation, include self-complementing split-GFP (Cabantous et al., 2005), split-superpositive GFP (Blakeley et al., 2012) and a tripartite split-GFP (Cabantous et al., 2013). It is important to note that the split proteins used in PCA do not reconstitute by way of their affinity for one another or their conformation, but solely upon interaction of their fused moieties. The study of split proteins as part of this project concentrates on the influence of their conformation on ligation.

As discussed earlier the balance between the size of the labelling components and their selectivity must be carefully considered. Split proteins would provide the specificity required, and would provide a regioselective template for a proximity-assisted chemical reaction to join the peptides and label the protein. Many of the split proteins detailed in Table 1.2 have small acceptor fragments that are potentially ideal for covalent attachment to the target protein, those investigated as part of this project are highlighted.

Protein	**K_d (nM)**	***N*-terminal peptide length (residues)**	***C*-terminal peptide length (residues)**
RNase A (Dwyer et al., 2001)	599	20	103
Cytochrome C (Hantgan and Taniuchi, 1977)	300	53	51
Thioredoxin (Holmgren, 1972)	6000	37	71
Chymotrypsin inhibitor 2 (de Prat Gay et al., 1994)	40	39	25
Monellin (Xue et al., 2004)	40	45	50
Protein G B1 (Kobayashi et al., 1995)	9000	41	15
Split ubiquitin (Müller and Johnsson, 2008)	7000	34	41
Dihydrofolate reductase (Smith and Matthews, 2001)	40	86	73
Synechocytis sp. *(Ssp)* DnaE (Shi and Muir, 2005, Wu et al., 1998)	36 ± 12	123	36
Calbindin D_{9k} (Berggård et al., 2001)	0.003	43	32

Table 1.2: A comparison of split protein peptide affinities and lengths

Adapted from Carey et al, 2007.

1.2.2 Choice of split proteins

The work detailed in the main results chapters is based on experiments involving bovine ribonuclease A (RNase A), chymotrypsin inhibitor 2 (CI2) and a calbindin D_{9k} variant. Each of these split proteins and the reason for their choice is detailed in the remainder of this chapter and in the introduction for their individual results chapters. Briefly, RNase A was the first demonstration of protein reconstitution (Richards, 1958, Richards and Vithayathil, 1959) and has been extensively studied, providing a thorough profiling in the literature for us to draw upon. Like RNase A, CI2 is a classic protein for folding and structural studies and its sequence too has been optimised (de Prat Gay and Fersht, 1994). In addition to the large amounts of previously published work, both of these protein systems were demonstrated to perform conformationally-assisted ligation (Beligere and Dawson, 1999). After investigating both of these systems, we also turned to a calbindin D_{9k} variant known in the literature for a strong affinity between its peptide fragments (Berggård et al., 2001) to attempt several conformationally-assisted chemical reactions.

1.2.2.1 RNase A

The first split protein system under investigation in this project was RNase A. In 1958 Frederic Richards demonstrated that subtilisin digestion of the 124 mer protein specifically produced two fragments of 20 and 104 residues termed the S-peptide and S-protein, ‘S’ denoting subtilisin digestion, respectively (Figure 1.1). This was the first demonstrated example in the literature of reconstitution, showing that the S-peptide and S-protein interact non-covalently to form the complex RNase S to regenerate ribonuclease activity (Richards, 1958, Richards and Vithayathil, 1959). Subsequent

work involved identifying those residues necessary for reconstitution, identifying the fewest residues required and those with better binding affinity (Potts et al., 1962, Dwyer et al., 2001, Shoemaker et al., 1987). It was found that just the first fifteen residues of the S-peptide were required for successful interaction with the S-protein, notably of the importance of the Asp-Ser dipeptide at positions 14 and 15 respectively (Potts et al., 1963). This was supported by the use of this 15-mer in the S-tag protein purification system as a carrier in fusion proteins with 20 fmol sensitivity of detection (Raines et al., 2000). The subtilisin digestion-liberated fragments were also shown to ligate with conformational assistance, using the structure of the two components under folding conditions to form a native amide bond at the native Ala-Ser site (Beligere and Dawson, 1999).

RNase A is investigated as the proof of principle system for this work, making use of the extensive background literature and optimisation. Although the tag acceptor peptide, the S-protein, is potentially too large to feasibly be used in the final labelling approach, it is commercially available and the S-peptide fragment can be prepared synthetically. Once suitable conditions have been established in this system we can then turn our attention to other split protein candidates.

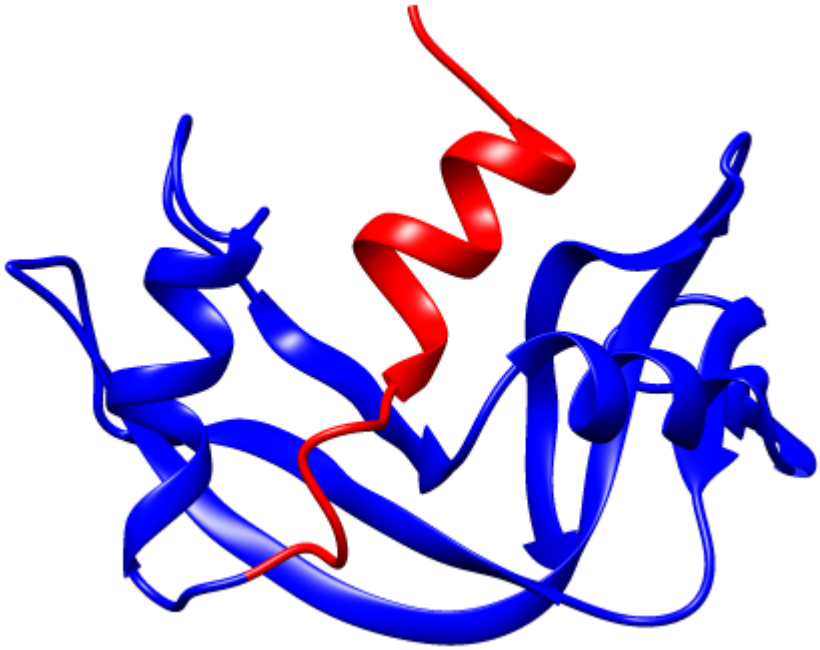


Figure 1.2: Structure of RNase A

3D structure of bovine RNase A (PDB 2AAS) as determined by NMR spectroscopy (Santoro et al., 1993). The S-peptide is coloured red and the S-protein coloured blue.

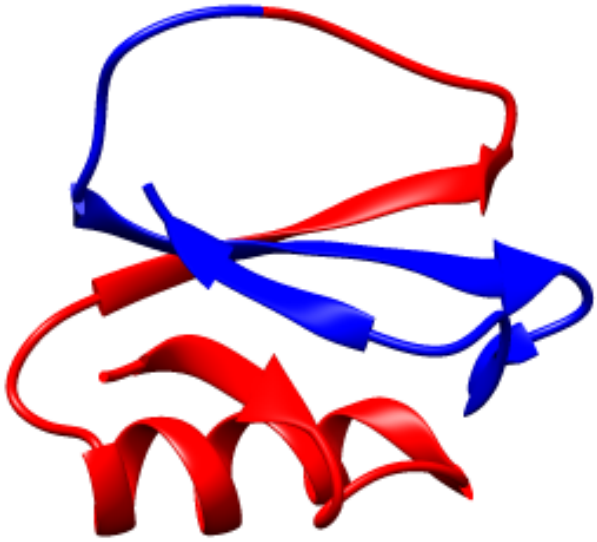


Figure 1.3: Structure of CI2

3D structure of barley CI2 (PDB 3CI2) as determined by NMR spectroscopy (Ludvigsen et al., 1991). CI2(20-59) is coloured red and CI2(60-83) is coloured blue.

1.2.2.2 CI2

CI2 was also chosen as a model split protein to study. As detailed in Table 1.2 its acceptor peptide is just 25 residues, preferable to the 104 residues in RNase A for potential use in the proposed method, and contains no Cys residues. It is a serine proteinase inhibitor consisting of 83 amino acid residues (Svendsen et al., 1980). The truncated protein was used in classic folding studies by the Fersht group (de Prat Gay et al., 1994, de Prat Gay and Fersht, 1994). This truncated form consists of 64 residues and was cleaved using cyanogen bromide (CNBr), selectively cleaving the peptide bond adjacent to the Met residue (Gross and Witkop, 1962), to generate CI2(1-39) and CI2(40-64) (Figure 1.2). The residue at position 40 is methionine, the only one in the sequence. The non-covalent re-association of CI2(1-39) and CI2(40-64) was demonstrated using circular dichroism (CD), fluorescence experiments and the inhibition of chymotrypsin (de Prat Gay and Fersht, 1994).

CI2 has been the subject of a number of folding studies with the component peptides generated by CNBr cleavage. The first four residues of the α–helix in CI2(20-59) play an important role in its stabilization (Elmasry and Fersht, 1994, de Prat Gay and Fersht, 1994) with the secondary and tertiary structure forming together during their association (de Prat Gay and Fersht, 1994). Conformationally-assisted ligation with these peptides was demonstrated, using the structure of the peptides under folding conditions to assist ligation (Beligere and Dawson, 1999). Remarkably, ligation was observed to take place not just with a Cys at the *N*-terminus of CI2(40-64) but with the native Met too (see Chapter 5).

1.2.2.3 Calbindin D_{9k}

The calbindin D_{9k} split protein was also chosen (Figure 1.3). It is a member of the calmodulin superfamily of proteins, calcium-binding proteins that share a common EF-hand motif. An EF-hand is a helix-loop-helix calcium binding subdomain, and is named after the E and F helices from the *C*-terminal subdomain of parvalbumin (Kretsinger and Nockolds, 1973). Calmodulin (Chattopadhyaya et al., 1992) is probably the best known member of this family but other members include the aforementioned parvalbumin, calbindin D_{28k} (Kojetin et al., 2006) and calbindin D_{9k} (Svensson et al., 1992). The $_{9k}$ and $_{28k}$ descriptions refer to the molecular weight, around 9 kDa and 28 kDa respectively. Calbindin D_{9k} is one of the smallest members of this family, is mainly expressed in mammalian enterocytes (Wasserman et al., 1992), and has two EF-hand subdomains that associate to form a 4-helix bundle. The first study of the individual EF-hands of calbindin D_{9k} used CNBr to cleave at a Met introduced in place of Pro43 in the connecting loop region (Finn et al., 1992). It is during this work that the Ser adjacent to the mutated Met43 residue and the associated problems are addressed. Cleavage with CNBr can generate a homoserine lactone from the iminolactone. A side reaction can take place if a Ser residue is adjacent, where the side chain of the Ser residue can attack the iminolactone to form a homoserine and an amide bond (Schroeder et al., 1969, Kaiser and Metzka, 1999, Corradin and Harbury, 1974, Proudfoot et al., 1989). This particular paper discusses methods to bypass this side reaction, and to reduce incomplete cleavage. The EF-hand peptides maintained structure after treatment with CNBr, demonstrating that the main contributor to their structure is the non-covalent forces between the peptides rather than the covalent bond between them. Stabilization of the peptides by calcium binding indicated that calcium coordination induces the preferred conformation for dimerization (Finn et al., 1992). Each EF-hand binds one

calcium ion, electrostatic interactions playing a key role in calcium coordination. In particular studies have shown that negatively charged residues at the binding site help (Linse et al., 1988), and mutants where surface carboxylates are replaced with amides demonstrated the role of protein surface charges (Linse et al., 1991). Fragment complementation studies involving the cleavage of calbindin D_{9k} between its two EF-hand subdomains, EF1 and EF2, established a K_d of 3 pM (Berggård et al., 2001). As such we were keen to investigate the potential of this protein for the proposed experiments (see Chapter 6).

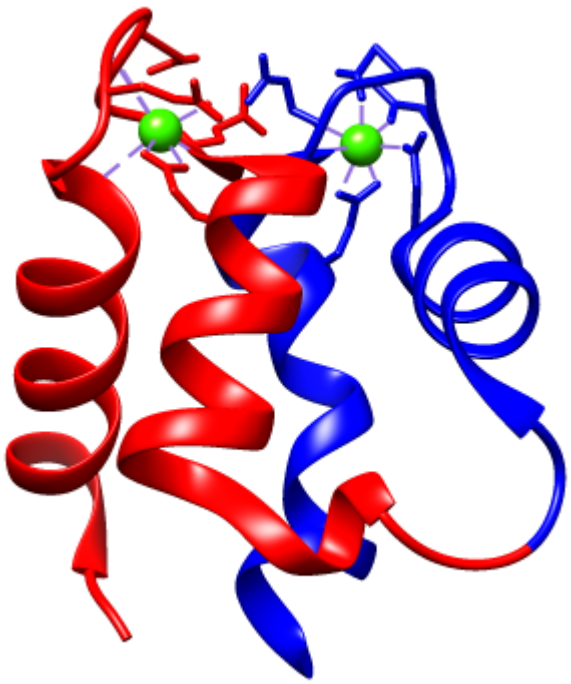


Figure 1.4: Structure of calbindin D_{9k}

Structure of calbindin D_{9k} as determined by X-ray crystallography (Svensson et al., 1992). EF1 is coloured red and EF2 is coloured blue, and the two coordinated calcium ions are shown in green.

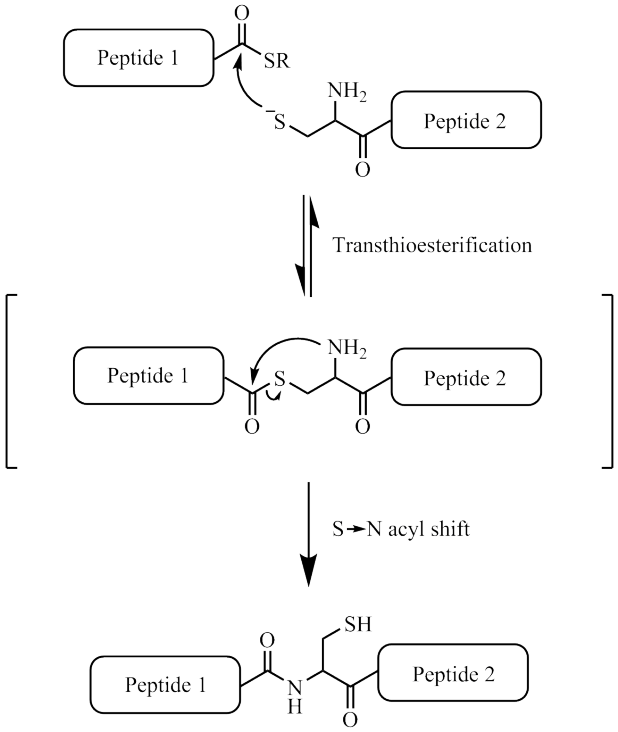
1.3 The development of conformationally-assisted chemical reactions

Several of the labelling methods discussed earlier in this chapter are based upon proximity-induced chemical reactions. In the proposed method the split protein would provide the proximity, so to establish the suitability of the reaction site in each of the chosen split protein systems the first experiments involved ligating the peptide at this site. Though these were conformationally-assisted ligations we must begin by first introducing the technique from which this form of ligation stemmed, and that is native chemical ligation.

1.3.1 Native chemical ligation

Native chemical ligation (NCL) involves the chemoselective coupling of two unprotected peptides, one with an *N*-terminal cysteine and the other with a *C*-terminal thioester, to form a native amide bond (Dawson et al., 1994) (Scheme 1.3). It has proven to be a very powerful technique providing synthetic access to proteins too large for solid phase peptide synthesis (SPPS) (Kent, 2009). Thioester peptide ligation to *N*-terminal cysteine peptides has been successfully demonstrated with most residues at the *C*-terminus, with only β–branched (Ile, Thr and Val) and Pro ligating slowly (Hackeng et al., 1999). The ability of split proteins that perform NCL to also undergo conformationally-assisted ligation was investigated during later stages of the project. Part of the generality of NCL is the use of denaturing conditions to overcome the poor solubility and aggregation properties of many peptides that can form structure in aqueous solutions, whereas conformationally-assisted ligation exploits peptide structure under folding conditions. It is a form of ligation that uses the structure and conformation

possessed by peptides to encourage ligation to occur, meaning that under folding conditions the *C*-terminus of one peptide will be in close proximity to the *N*-terminus of the other.



Scheme 1.3: Proposed mechanism for native chemical ligation

Two unprotected peptides, one with a *C*-terminal thioester and the other with an *N*-terminal Cys, couple chemoselectively and form a native amide bond at neutral pH. The first step involves the nucleophilic attack of the α–carboxyl group of peptide 1 by the Cys side chain of peptide 2, forming a thioester-linked intermediate. This intermediate then rearranges and through an *S*→*N* acyl shift the peptides are linked by a native amide bond.

1.3.2 Conformationally-assisted ligation

Conformationally-directed chemical reactions have long been a subject of interest. They differ from NCL in that they do not necessarily require a Cys at the reaction site, but rely on the conformation of the components to bring the reacting site termini into close proximity. This proximity raises the effective concentration of the reactive components (Way, 2000), and induces ligation.

Early work with cytochrome *c* showed that component fragments are able to form complexes echoing the conformation and, to an extent, the activity of the native protein, particularly if the fragments are sequential (Wallace and Proudfoot, 1987, Proudfoot et al., 1989, Wallace et al., 1986). The reaction is catalysed by the proximity of the peptides that increases the effective concentration of the components at the reactive site (Wallace et al., 1986), a mechanism upon which most semisyntheses of cytochrome *c* depended. The majority of syntheses of cytochrome *c* rely on CNBr cleavage, an important step due to the generation of few fragments by the low frequency of Met residues in proteins. Early work exploring this developed conditions for cleavage at the accessible Met65 rather than the heme-coordinating Met80 (Corradin and Harbury, 1970), and showed that a mixture of the fragments generated formed a reconstituted 1:1 complex in aqueous solution (Corradin and Harbury, 1971). Of particular interest was the homoserine lactone coupling method, whereby the homoserine lactone formed from CNBr treatment can spontaneously re-form the peptide bond. More recent work has exploited the use of CNBr cleavage to generate the *N*-terminal homoserine lactone peptide for the semisynthesis of cytochrome *c* to incorporate carbon-deuterium bonds for time-resolved studies on labelling and folding measurements (Sagle et al., 2004, Sagle et al., 2006).

Templated synthesis combines chemical ligation with self-assembling peptides, early work in this area using coiled-coils to template a ligation between a *C*-terminal thioester peptide and an *N*-terminal Cys peptide (Severin et al., 1997, Yao et al., 1997). Ligation at a Cys-free site has been effected using leucine-zipper sequences in a self-assembling fiber system containing peptides (Ryadnov and Woolfson, 2007). Rational design of coiled-coils has established systems with dissociation constants in the subnanomolar range that are being used for exploration of protein-protein interactions (Thomas et al., 2013).

Coiled-coils consist of heptad repeats of hydrophobic and polar residues, and when folded into an α-helix create a hydrophobic interface through which they interact and assemble (Crick, 1953, Pauling and Corey, 1953, Lupas, 1996). We chose not to use coiled-coils in the labelling method to avoid the potential risks associated with oligomerization, but additionally we wanted a specific ligation at the *N*-terminus of the expressed peptide and not in the middle of the sequence as with coiled-coils. As a result we focused more on the conformationally-assisted ligation of peptides at a Cys-free site under folding conditions, as with the conformationally-assisted ligation of RNase A and CI2 (Beligere and Dawson, 1999) discussed earlier in this chapter and later in Chapters 4 and 5.

1.3.3 Click chemistry

When conformationally-assisted ligation was observed between split protein peptide fragments it was thought that perhaps a similar approach could be applied to the development of a reaction usually requiring a catalyst to proceed. One branch of click chemistry in particular, Huisgen 1,3-dipolar cycloaddition, has been the subject of a

number of studies looking to obviate the need for a copper (Cu) I catalyst (Jewett and Bertozzi, 2010), and so we wondered if the calbindin D_{9k} peptides could ligate in this manner.

The term "click chemistry" was first applied by Sharpless and colleagues, and covers a range of reactions that aim to join together small units *via* heteroatom links (Kolb et al., 2001). For a reaction to be classed as following click chemistry it must adhere to a number of criteria including simple reaction conditions, high yields, wide scope, benign/no solvent, readily available starting materials and stability under physiological conditions. They are described as "spring-loaded" reactions due to their selectivity for and rapid reaction toward a specific product. One area of click chemistry covers cycloaddition reactions, probably the most popular of which is Huisgen dipolar cycloaddition (Huisgen, 1963). Original Huisgen cycloaddition between an azide and a terminal alkyne to form a triazole required elevated temperatures in order to proceed, until two separate groups performed ligation at room temperature using a Cu(I) catalyst (Rostovtsev et al., 2002, Tornoe et al., 2002). Such a reaction, referred to as Cu(I)-catalysed azide-alkyne cycloaddition (CuAAC), solely generates 1,4-distributed 1,2,3-triazoles (Scheme 1.4), though such a metal catalyst is toxic to the cell (Baskin et al., 2007).

Efforts to achieve click chemistry in the absence of copper include the use of ring strain with strained alkynes (Agard et al., 2004, Li et al., 2004). However prior to the inclusion of copper catalysts, work had been conducted whereby cycloaddition could be accelerated by the templating of the reactants. The formation of a triazole is accelerated by the addition of cucurbituril, a nonadecacyclic caged compound (Freeman et al., 1981), that binds the peptide azide and alkyne to orientate them favourably (Mock et al., 1983, Mock et al., 1989). This rate improvement relies upon the formation of this

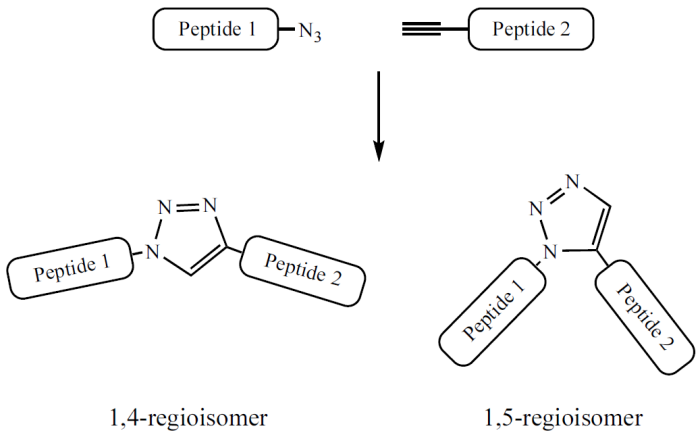
transient complex, and follows the Pauling principle of catalysis: the binding capacity of cucurbituril favours the reaction transition state rather than the substrates, with the complementarity of cucurbituril greater for this intermediate.

A proximity-induced thermal 1,3-dipolar cycloaddition forming a 1,2,3-triazole product has also been used to identify acetylcholinesterase (AChE) inhibitors (Lewis et al., 2002), a form of target-guided click chemistry where recognition results from the spatial orientation of the peptide reactive components. The binding pocket of AChE is a narrow cleft into which ligands containing an azide or alkyne were introduced, the confines of the cleft encouraging the reaction to take place by initiating favourable orientation of the azide and alkyne. This method is known as small molecule *in situ* click chemistry (SISC) and has also been used to identify inhibitors in a number of other systems (Mamidyala and Finn, 2010, Millward et al., 2013). A recent report on the search for histone deacetylase (HDAC) inhibitors found that the presence of trace amounts of Cu associate with the enzyme in its active site increased the rate of triazole formation, combining proximity-assisted and Cu-catalyzed AAC (Suzuki et al., 2010). Studies such as these confirm that these components are able to function with stability in water, are easy to incorporate into the system of choice, and compatible with physiological conditions (Saxon, 2000, Kiick et al., 2001, Baskin et al., 2007).

A similar form of proximity-assisted click chemistry is iterative peptide *in situ* click chemistry (IPISC), focusing on the identification of ligands that bind target proteins with antibody-like affinity. The aim is to combine the affinity and specificity seen in monoclonal antibodies with the synthetic accessibility of peptides. The binding of a primary ligand to the target protein identifies a secondary ligand, and the peptides bind to form a triazole-linked biligand. The biligand can then be used to identify a

tertiary ligand, and so on. This method was first used to generate a triligand for carbonic anhydrase II (Agnew et al., 2009).

In situ click chemistry, with examples such as the AChE work discussed earlier, preceded CuAAC and has been used in the design of synthetic receptors that can regulate specific protein-protein interactions, such as the Grb2-SH2 domain mimic that was shown to induce cytotoxicity and inhibition of tumour growth in cancer cells (Tanaka et al., 2012). The use of low concentrations of Cu(I) provide slow 1,3-dipolar cycloaddition to reduce background reactions. In this instance the anchor is resin-bound bis-Lys that binds to phosphate, modified to contain two azides, and the target is a phosphorylated cyclic peptide selective for the Grb2-SH2 domain. The azides form triazole linkages with members of an alkyne-containing tetrapeptide library, and specifically binds to Grb2-SH2 interacting proteins both *in vitro* and *in vivo*.



Scheme 1.4: Huisgen 1,3-dipolar cycloaddition products

A peptide azide and a peptide alkyne can ligate to form the 1,2,3-triazole product. This product can take the form of either the 1,4- or the 1,5-regioisomer, but solely the 1,4-regioisomer if a Cu(I) catalyst is used (Tornoe et al., 2002, Rostovtsev et al., 2002).

2. Materials and methods

2.1 Reagents and chemicals

All chemicals and reagents purchased from Sigma Aldrich with the exception of the following.

All L-amino acids and Fmoc-Asp(OtBu)-Ser(ψMe,Mepro)-*OH* were purchased from Novabiochem, in addition to Fmoc-Gly-4-sulfamylbutyryl rink amide AM, Fmoc-Ala-sulfamylbutyryl sieber amide, Fmoc-PEG biotin NovaTag and azido NovaTag resins. PL rink and Fmoc-Gly-Wang resins were purchased from Agilent Technologies. Tris(2-carboxyethyl)phosphine hydrochloride (TCEP-HCl) and guanidine hydrochloride (Gdn.HCl) were purchased from Thermo Scientific. 1-hydroxybenzotriazole (HOBT) was purchased from CS Bio. Trifluoroacetic acid (TFA), piperidine and acetonitrile (AcN) were purchased from Romil. *N,N*-Dimethylformamide (DMF) in bottles was purchased from Severn Biotech Ltd., and in 25 L drums from AGTC Bioproducts Ltd. Diethyl ether, dichloromethane (DCM) and ethylenediaminetetraacetic acid (EDTA) were purchased from Fisher Scientific. Calcium chloride was purchased from VWR International. *N*α,*N*ε-bis-Boc-L-Lys-*OH* was purchased from Santa Cruz Biotechnology.

The LAPAG-MPAL thioester peptide of section 6.4 was synthesized and generously donated by Richard Raz.

2.2 Buffers and cocktails

All buffers are specified in figure legends. Buffers and cocktails in the main text adhere to the following recipes, changes or additive specified where included.

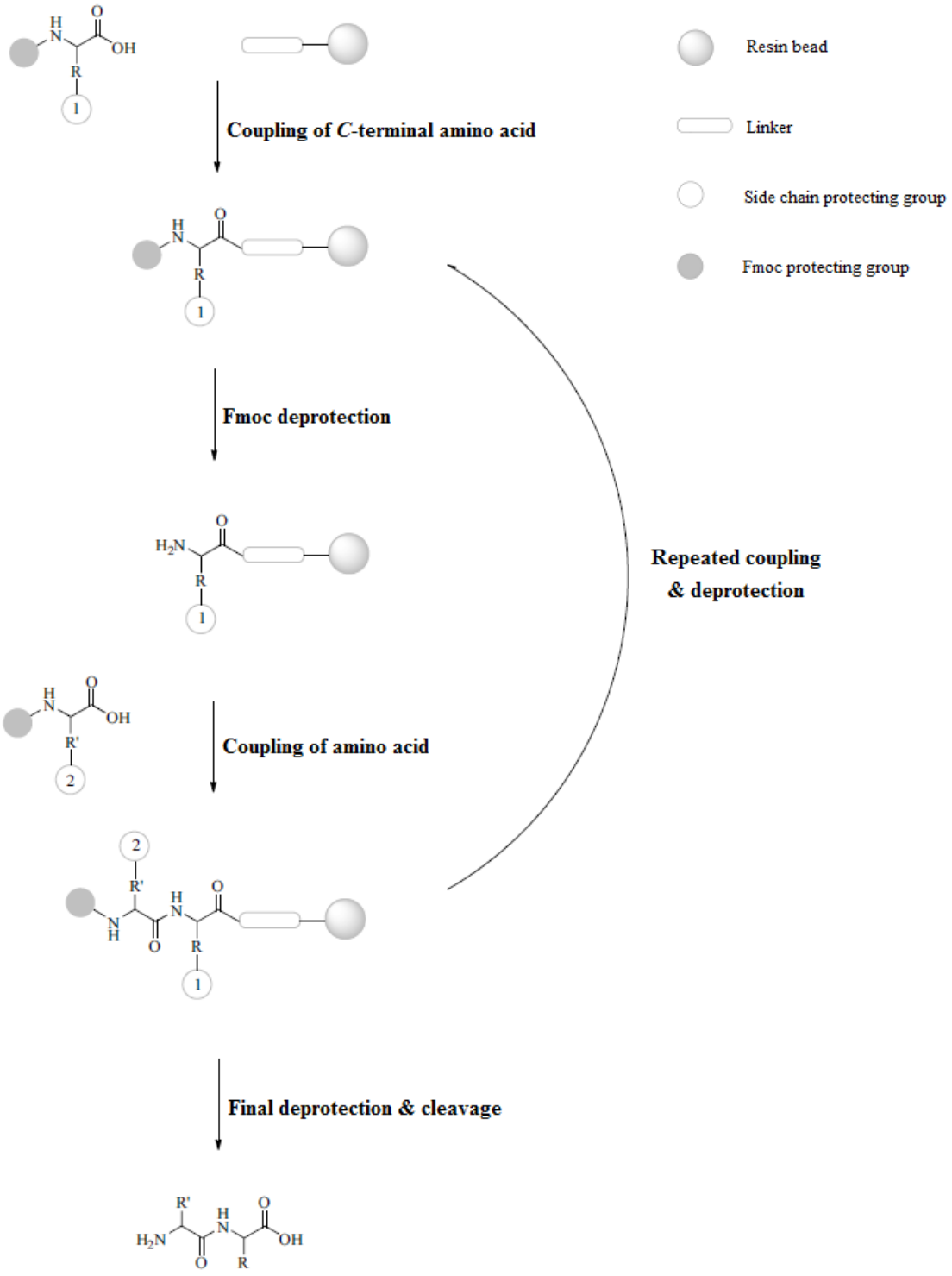
- HPLC buffer A: TFA (0.1 %) in H_2O.
- HPLC buffer B: TFA (0.1 %), AcN (90 %) in H_2O.
- LCMS buffer A: formic acid (0.1 %) in H_2O.
- LCMS buffer B: AcN.
- Denaturing buffer: Gdn.HCl (6 M), sodium phosphate (200 mM), EDTA (2 mM), pH 7.5.
- Non-denaturing buffer: sodium phosphate (100 mM), EDTA (2 mM), pH 7.5.
- Calcium non-denaturing buffer: sodium phosphate (100 mM), calcium chloride (2 mM), pH 7.5.
- BLI/CD calcium buffer: HEPES (10 mM), Tween (0.005 %), calcium chloride (2 mM), pH 7.0.
- CD EDTA buffer: HEPES (10 mM), Tween (0.005 %), EDTA (2 mM), pH 7.0.
- General cleavage cocktail: TFA (5 ml), H_2O (250 µl) and triethylsilane (TES, 250 µl). 1,2-ethanedithiol (EDT, 100 µl) included for Met/Cys-containing peptides.

2.3 Preparation of peptides

2.3.1 9-Fluorenylmethoxycarbonyl solid phase peptide synthesis

In 9-Fluorenylmethoxycarbonyl solid phase peptide synthesis (Fmoc-SPPS) the α–amino of each residue is protected by the base-labile Fmoc protecting group, unlike the acid-labile t-butoxycarbonyl (Boc) group of Boc-SPPS. The side chain protecting groups and resin linker are acid labile, and thus can be cleaved simultaneously. The use of milder conditions during these steps makes the Fmoc method more readily accessible than the Boc method: the HF required in Boc-SPPS is highly dangerous and handling requires specialist equipment.

The peptide chain is built off the solid support *via* a linker that provides a cleavage site and determines the functionality of the *C*-terminus. Each amino acid has side chain protection to prevent unwanted side reactions between potentially reactive groups, and each amino acid is introduced with an Nα-protecting group to ensure coupling of one residue at a time. Depending on the synthesizer used, the amino acid and coupling reagents are either pre-mixed or are mixed in the reaction vessel itself. Simply, a single cycle in Fmoc-SPPS will begin with deprotection of the Nα-protecting group, the Fmoc group. With the *N*-terminus of the amino acid now free the next amino acid building block, activated by the coupling reagents, is added. With sufficient wash steps between each process the cycle is ready to begin again, until the synthesis is complete and the peptide is cleaved (Scheme 2.1). Microwave-assisted Fmoc-SPPS uses microwave radiation and elevated temperatures during the coupling and deprotection steps to reduce reaction times (Bacsa et al., 2006).



Scheme 2.1: Overview of Fmoc-SPPS

The first step is to couple the *C*-terminal amino acid to a resin bead with linker. Following deprotection of the Fmoc group of this amino acid, the next amino acid is coupled. This cycle of Fmoc deprotection and subsequent amino acid coupling is repeated for the rest of the sequence. Final deprotection of the Fmoc group is followed by simultaneous cleavage of the peptide from the resin and cleavage of the side chain protecting groups, yielding the unprotected peptide target. R and R' = amino acid side chains.

Two common methods for testing successful deprotection or coupling are the Kaiser test (Kaiser et al., 1970) and the 2,4,6-trinitrobenzene-sulphonic acid (TNBS) test (Hancock and Battersby, 1976). The first is a reaction of ninhydrin with amines, and is most sensitive to primary amines. The second test involves adding mixtures of *N,N*-diisopropylethylamine (DIEA) and TNBS to resin beads washed with dichloromethane (DCM). The dibenzofulvene released during Fmoc cleavage is a strong chromophore and provides an additional avenue for monitoring deprotection by UV spectroscopy.

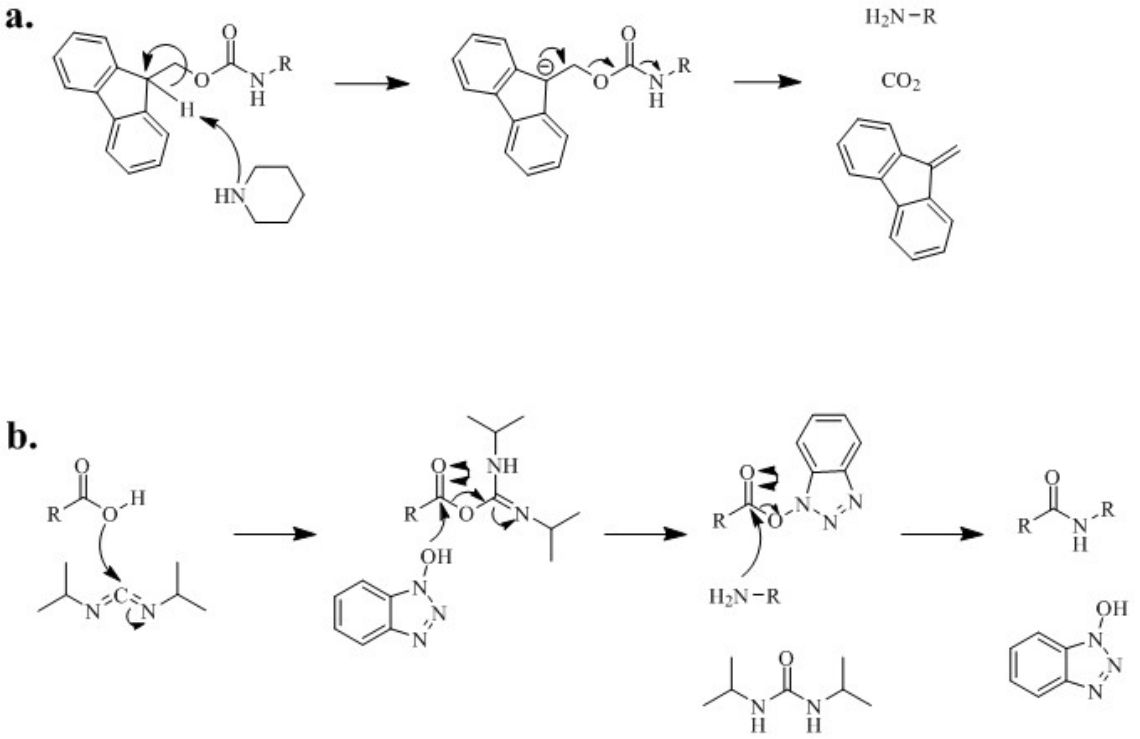
Upon completion of the synthesis the cleavage of both the peptide from the resin and the side chain-protecting groups occurs simultaneously when treated with trifluoroacetic acid (TFA). Scavengers are also required to prevent the side chain deprotection from reversing. Common scavengers include water, triethylsilane (TES) to quench carbocations and 1,2-ethanedithiol (EDT) for peptides containing Met or Cys whose deprotected side chains are acid-sensitive.

Fmoc-SPPS is not without its side reactions, which must be managed in order to achieve efficient peptide synthesis. One such problem encountered in this project is reduced solubility, resulting in aggregation that can affect the efficiency of both amino acid coupling and deprotection. This is the self-association of the peptide on-resin that often leaves the reaction site obstructed. One aid in minimizing aggregation is the introduction of pseudoprolines into a sequence. A pseudoproline is an oxazolidine derivative of either a Ser or Thr residue, being the *C*-terminal residue of a dipeptide (Wohr and Mutter, 1995). These building blocks introduce a ‘kink’ into the peptide backbone by inducing a *cis* amide bond, acting like Pro to disrupt regular structure and hydrogen-bonding. The TFA treatment at the end of the synthesis to simultaneously cleave the side chain-protecting groups and the peptide from the resin leads to ring opening and the leaving of a regular Ser or Thr at the site (Haack and Mutter, 1992).

Another solution is to add N-(2-hydroxy-4-methoxybenzyl) (Hmb)-protected amino acids into the synthesis, ideally every six or so residues (Johnson et al., 1993, Hyde et al., 1994, Simmonds, 1996), preceding the problematic area to disrupt the problematic hydrogen-bonding. Like pseudoprolines, these are easily cleaved at the end of the synthesis during the TFA treatment. Hmb protection is also particularly useful in the prevention of aspartimide formation, where the side chain of Asp or Asn can attack the amino group of the preceding amino acid following proton abstraction from the nitrogen under basic conditions. The main factors that can encourage aspartimide formation are the base used and the *C*-terminal amino acid, particularly when preceded by Gly, Ser or Thr (Bodanszky and Kwei, 1978, Nicolas et al., 1989). The use of the Hmb protecting group has proven very powerful in the prevention of aspartimide formation by providing backbone protection (Quibell et al., 1994, Offer et al., 1996).

For this work peptides were synthesized by Fmoc-SPPS using automated solid phase synthesis, either with (CEM Liberty 1) or without (CS Bio) microwave-assisted coupling and deprotection. For microwave-assisted syntheses deprotection of the Fmoc group was achieved using piperidine (20 %) in DMF (Scheme 2.2a), and subsequent amino acid activation/coupling using *N,N'*-diisopropylcarbodiimide (DIC, 0.8 M) in dimethyl sulfoxide (DMSO) and HOBT (0.5 M) in DMF (Scheme 2.2b). Amino acids (0.2 M) in DMF were coupled with a 10 min, 75 °C, 25 W microwave-assisted coupling. β–branched amino acids and their *N*-terminal residue were double coupled, to combat potential reduced coupling efficiency due to steric hindrance. An initial deprotection of 30 s at 75 °C and 25 W was followed by a 3 min, 75 °C, 25 W microwave-assisted deprotection.

Any alteration to these standard conditions is detailed in the appropriate section.



Scheme 2.2: Fmoc deprotection and coupling mechanisms

a. Fmoc deprotection. The acidic proton eliminates a dibenzofulvene during base treatment *via* an E1cB elimination, the two phenyl groups helping to stabilize the anion. Loss of CO_2 generates an amine, allowing for the coupling of the next amino acid building block.
b. DIC/HOBT coupling. The carboxyl group of the Nα-protected amino acids must be activated in order to couple to the resin-bound deprotected peptide. DIC activates the amino acid by providing a better leaving group, and couples to form a transient O-acylurea. To minimise racemization HOBT is introduced and reacts to form a less reactive activated ester. This then couples to the preceding deprotected amino acid in the sequence, forming an amide bond.

2.3.2 Cleavage from resin and purification

Resin subjected to a DCM wash (5 x 10 ml) and dried under vacuum (ILMVAC GmbH) for weighing, and divided into two portions prior to cleavage. To each either the general or EDT-containing cleavage cocktail was added, depending on the presence of cysteine or methionine residues. Following a 1.5 h cleavage time the resin was filtered, the TFA sparged (N_2), and the peptide precipitated in Na-dried ether (40 ml). After cooling on dry ice for a minimum of 10 min the precipitated peptide was pelleted by centrifuging (4000 rpm for 10 min at 4 °C, eppendorf 5810 R centrifuge), and the ether decanted. Peptides were resuspended in either HPLC buffer A with minimal buffer B or, if solubility was more problematic, the denaturing buffer.

Samples were filtered prior to purification using a Whatman 13 mm disposable PVDF filter with 0.2 μm pore size. Buffer A was filtered (Nalgene 0.2 μm) and both buffers degassed (ILMVAC GmbH) for preparative high-performance liquid chromatography. Purification by preparative-HPLC (Waters 600E) with an appropriate linear gradient, monitoring at 214 nm using either a Grace Vydac RP-C18 (218TP1022 22 x 250 mm) with a flow rate of 15 ml min^{-1}, or an RP-C18 (218TP1010) with a flow rate of 5 ml min^{-1}. Sample-containing fractions were pooled according to purity and freeze-dried.

2.3.3 Peptide characterization

2.3.3.1 Matrix assisted laser desorption/ionization time-of-flight mass spectrometry

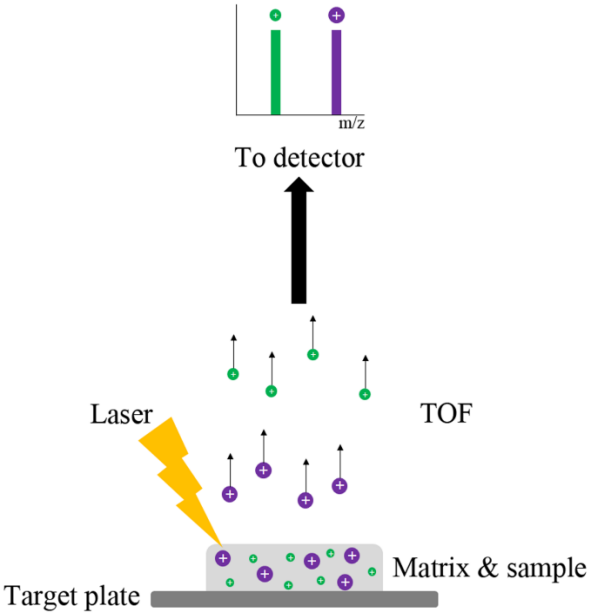
Matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was developed by Tanaka (Tanaka et al., 1988), Hillenkamp and

Karas (Karas et al., 1985), and involves ionization of the sample by a laser before collecting the resulting ionized molecules. Firstly the sample is mixed with a matrix that transfers the ionization energy from the laser to the sample, with around 1 µl added to the target plate and allowed to co-crystallize. There are a variety of matrix substances depending on the type of sample in question, due to differences in the energy needed for ionization along with sample stability. The sample/matrix mix is then shot with a laser to generate ions, where desorption by the laser beam results in ablation of the top layer of the matrix/sample mix and subsequent ionization of the sample. Ionization of neutral molecules [M] typically generates $[M+H]^+$ with a proton added, $[M-H]^-$ with a proton removed, $[M+nH]^{n+}$ with multiply charged ions, $[M+Na]^+$ with a sodium ion added and $[M+K\}^+$ with a potassium ion added. These ions must then be detected and analysed (Scheme 2.4).

A TOF mass spectrometer separates these charged ions according to their flight times. The linear mode of analysis gives limited mass resolution and so the homogeneity of the sample/matrix mix is crucial for optimising the balance between resolution and spatial spread, whereas reflectron mode improves resolution by efficient ion focusing. The reflectron has an 'ion mirror' that reflects ions using an electric field, doubling the flight path of the ionized molecules and thus improving mass accuracy. It diminishes the spread of the flight times of ionized molecules with the same mass-to-charge ratio (m/z).

Prior to analysing a sample an external calibrant is typically used to remove any error generated from moving the stage in and out when loading the target plate. The calibrant is a known mixture of proteins or peptides corresponding to a list in the analysis software, allowing for peaks to be assigned to the correct mass if necessary. This calibration can then be applied to sample spectra.

MALDI-TOF MS (Bruker microflex) used α-cyano-4-hydroxycinnamic acid (CHCA) matrix (10 mg ml^{-1} in H_2O:AcN:TFA 50:50:0.1), generally with ion positive reflector mode for peptides with m/z < 4000 and ion positive linear mode for > 4000.



Scheme 2.3: Outline of MALDI-TOF MS

The sample is mixed with the appropriate matrix and dotted onto the target plate. A laser ionizes the sample molecules which are separated by their flights times in the TOF MS, with smaller molecules (green) flying faster than larger molecules (purple), before arriving at the detector.

2.3.3.2 Liquid chromatography-mass spectrometry

Liquid chromatography-mass spectrometry (LCMS) combines the sample separation of HPLC with the mass analysis of MS, most often using electrospray ionisation (ESI) to link the two processes. One advantage that this method holds over MALDI is that it can produce multiply charged ions, therefore extending the mass range of the analyser.

The sample is introduced to the mobile phase and pressurized for column separation, with pumps controlling the composition and flow rate. Desolvation during ESI gives a fine aerosol that is dispersed by electrospray, and non-volatile molecules with multiple charges are now in the gas phase. For the specific equipment used this method can detect peptides and proteins up to 100 kDa, and from as low as 150 Da. The detector in this model is a single quadrupole mass analyser, and can detect both positive and negative ions. This mass analyser sorts ions based on their m/z, their separation based upon the stability of their trajectories along the length of the four rods. A radio frequency voltage is applied between the pairs of rods, and a direct current voltage is superimposed. Only ions with a certain m/z will travel the length and reach the detector under a specific ratio of these applied voltages, and others will have unstable trajectories and hit the rods. This allows for selection of a specific m/z but also for monitoring a range by varying the voltages.

LCMS (Shimadzu LCMS-2010 EV) used a Shim-pack XR-ODS II 2.0 mmi.d x 75 mm column with a linear gradient of 0 – 95 % B in 10 min and a flow rate of 0.5 ml min^{-1}, monitoring at 214 nm. LCMS was used for peptide characterization during peptide preparation.

2.3.3.3 High performance liquid chromatography

High performance liquid chromatography (HPLC) is a technique for separating and identifying components of a mixture. It can also be used to purify a sample, for example following SPPS, by separating the components and delivering them to a fraction collector.

A pressurized system controls solvent flow and is used to move the sample through a column to separate, separation depending on the type of column used and the interaction of the molecules with the column material. The sampler delivers the sample from the injection loop into the 'mobile phase', a mixture of solvents, the flow rate and composition of which is controlled by the pump. Once the sample has passed through the column the detector relays the information gathered into a signal corresponding to the separation of the sample molecules.

Reverse phase-HPLC (RP-HPLC) separates molecules based on their hydrophobicity, using a non-polar stationary phase and a polar mobile phase. The most common stationary phase is probably silica modified such that the surface of the particles is coated with hydrophobic alkyl chains of varying lengths, such as the C18 and C4 columns used during this project. The mobile phase can be of constant (isocratic flow) or varied (gradient flow) composition. In gradient flow 'solvent A' refers to the aqueous component and 'solvent B' refers to the organic solvent, typically water and acetonitrile respectively. Less polar molecules will have longer retention times, and retention times can be controlled by the composition of the mobile phase. Detectors typically monitor at 214 nm to detect the absorbance of the peptide backbone, unless using a diode array detector (DAD) that monitors a range of wavelengths.

During this project analytical-HPLC (Agilent 1200 series, and VWR Hitachi ELITE LaChrom) generally used linear gradients with a variety of columns, monitoring at 214 nm for determination of peptide purity and retention time. Columns specified in appropriate sections, and all figure legends.

For the duration of this work, and specified in figure legends, the following columns were used:

- Grace Vydac RP-C18 (218TP 5µ 150 mm x 4.6 mm)
- Phenomenex RP-C18 (Gemini 3µ 150 mm x 4.6 mm)
- Grace Vydac RP-C4 (214TP 5µ 150 mm x 4.6 mm)
- Chromolith RP-8 column (Merck RP-8e 4.6 mm x 100 mm)

2.3.3.4 Yield calculation

Peptide yields were calculated from the weight of final dried peptide, its mass including that of the CF_3COO^- counter ion, in relation to either the scale of the synthesis based upon the loading of the resin or the amount of starting materials used.

2.4 Biophysical characterization

2.4.1 Ultraviolet/visible spectroscopy

Ultraviolet/visible (UV/VIS) spectroscopy is a form of absorbance spectroscopy, measuring transitions from the ground state to the excited state in molecules containing π electrons that can absorb light in this spectrum. Using a UV/VIS spectrophotometer light passes from the radiation source through the monochromator or prism to separate the wavelengths, and any light that passes through is measured by the detector. The intensity of light transmitted through the sample-containing cuvette (I) is compared to the intensity before it passes through (I_0). This ratio gives the absorbance (A) by the equation:

$$A = \log_{10}(I_0/I)$$

One use of this technique is the determination of a sample's concentration in solution. A number of biophysical techniques used in this project require accurate concentrations, notably circular dichroism and bio-layer interferometry. Molecules with chromophores have specific absorbance maxima at a set wavelength. These maxima (λ_{max}) can be plotted as an absorbance spectrum and used to identify concentration. A compound-specific extinction coefficient (Ɛ) for a set wavelength can be used for calculation using the Beer-Lambert law:

$$A = Ɛ \text{ x } c \text{ x } l$$

where A is the absorbance at 280 nm (A_{280}), Ɛ is the extinction coefficient (5560 M^{-1} cm^{-1} for each Trp, 1200 $M^{-1}cm^{-1}$ for each Tyr), c is the concentration and l is the path length. This states that the absorbance of chromophores such as Trp and Tyr is directly proportional to their concentration and the path length.

Another form of UV/VIS spectroscopy uses a NanoDrop instrument. This technique is particularly useful when dealing with small volumes, such as the ligations discussed in later chapters, as no cuvette is used and so there is no need for a minimum volume. µl amounts are pipetted directly onto the instrument's pedestal, and when the arm is lowered to make contact with the droplet a sample column forms. The pedestal then moves to adjust to the ideal path length, and the measurement is conducted in much the same way as described above. Advantages include easy sample preparation, quick measurements and requiring just µl volumes of sample.

UV/VIS spectroscopy (Jasco V-550 spectrophotometer) was employed to determine peptide concentration. The model used has a double beam optical system with a single monochromator, with deuterium and halogen lamps and a photomultiplier R928 detector. Absorbance at 280 nm (A_{280}) was measured using a cuvette with 10 mm light path (Hellma Suprasil quartz 105B-QS). Concentration was calculated using the Beer-Lambert law once a value for absorbance has been established.

NanoDrop2000 (Thermo Scientific) was used to determine ligation component concentrations, monitoring absorbance at 280 nm.

2.4.2 Fluorescence spectroscopy

Fluorescence spectroscopy is a form of electromagnetic spectroscopy that analyses the fluorescence of a sample, where excitation of the electrons of aromatic amino acids by UV light results in light emission. UV light passes through a filter or a monochromator to select the wavelength before hitting the sample in solution in a cuvette with a determined path length. Fluorescence light is emitted by specific residues

and is passed through a second filter or monochromator prior to reaching a photomultiplier detector, usually at 90 ° to prevent the reflection of the incident light from reaching the detector. Most fluorescence emitted comes from Trp but there can also be emission due to Tyr and Phe residues. Fluorescence spectroscopy can be used to study protein folding, whereby the fluorescence emitted by these residues will decrease as the protein folds and they become buried.

Fluorescence data were collected (Varian Cary Eclipse fluorescence spectrophotometer) by monitoring absorbance at 345 nm and using a cuvette with 3 mm light path (Hellma Suprasil quartz 105.251-QS). The model used has a Xe lamp, monochromators with filters, Schwarzchild collection optics and photomultiplier tube detectors.

2.4.3 Bio-layer interferometry

Bio-layer interferometry (BLI) is an optical technique based upon the analysis of a white light interference pattern, and is a label-free method for monitoring biomolecular interactions. White light is reflected from two places, one being the interface between solvent and protein bound at the tip of the sensor and the other an internal reference point. As the partner peptide interacts with the bound peptide this layer of protein becomes thicker, causing a shift in the interference pattern that is monitored in real time (Scheme 2.5). Only peptides that are bound at the sensor tip contribute, therefore unbound molecules have no effect. It could thus be said that BLI essentially measures a change in mass.

BLI measures the equilibrium dissociation constant (K_d) from a ratio of kinetic constants. For a 1:1 reaction the time dependence of the response from the sensor (R) during the association phase, when the when the partner protein binds that at the tip, is described by the equation:

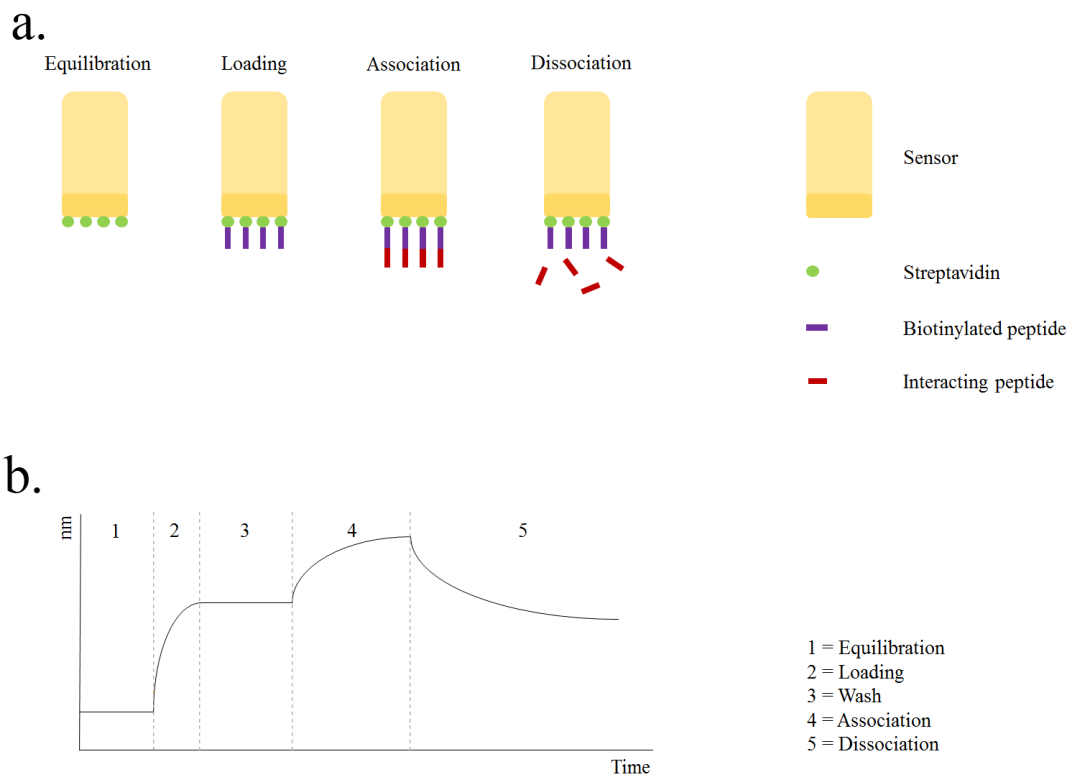
$$R = R_{eq}(1 - e^{-k_{obs}t})$$

where k_{obs} is a pseudo first-order rate constant (= $k_{on}[C] + k_{off}$, where k_{on} is the association rate constant, k_{off} is the dissociation rate constant and [C] is the concentration of the interacting protein) and R_{eq} is the equilibrium response. The dependence of k_{obs} on the concentration of the interacting protein for its bound partner allows for the determination of k_{on} from the slope of k_{obs} plotted against [C]. When the intercept is too close to zero the k_{off} value cannot be accurately determined. In such instances k_{off} can be calculated from the dissociation phase, where the time dependence of R is then described by the equation:

$$R = R_0 e^{-k_{off}t}$$

where R_0 is the biosensor response before the start of dissociation.

BLI was carried out using the Octet Red system (FortéBio). Streptavidin sensors (FortéBio SA Dip and Read biosensors for kinetics) were allowed to soak in the BLI calcium buffer of HEPES (10 mM), Tween (0.005 %) and calcium chloride (2 mM), pH 7.0 for ≥ 20 min prior to experimentation. 96 well plates included a row of buffer for washing, a row of the biotinylated peptide (including buffer blank wells as control) and rows of corresponding partner peptide at varying concentrations.



Scheme 2.4: Outline of an Octet experiment with a biotinylated peptide and streptavidin-coated sensors

a. Steps involved in typical Octet BLI experiment. Following equilibration of streptavidin-coated sensor in buffer, the biotinylated peptide is loaded onto it. When this is added to the interacting peptide the peptides associate, before their dissociation is then measured.

b. Example of an Octet experimental trace. 1: A streptavidin-coated sensor is allowed to equilibrate in buffer. 2: The sensor is then moved to the wells containing biotinylated peptide, for loading. 3: The loaded sensor is then replaced in the buffer, to wash away any unbound biotinylated peptide. 4: The loaded sensor is moved to the wells containing the interacting peptide, to monitor their association. 5: The loaded sensor with bound interacting peptide is then replaced into the buffer wells to wash away the interacting peptide, to monitor their dissociation.

2.4.4 Circular dichroism spectroscopy

Circular dichroism (CD) is the differential absorption of left- or right-handed circularly polarized light by optically active proteins, and is an effective method for monitoring conformational changes of a sample. Light can be polarized using filters or prisms to cause the electric field I to oscillate in a sinusoidal manner in one plane, either rotating clockwise or anticlockwise when viewed front-on. During a CD experiment using far-UV light to investigate protein secondary structure a sample is subjected to equal amounts of both left- and right-circularly polarized light alternately. One will be better absorbed than the other and this difference is measured. Data in millidegrees are converted to ΔA ($= A_{left} - A_{right}$) and then to mean residue ellipticity (MRE, deg cm^2 $dmol^{-1}$ res^{-1}), normalized for the concentration of peptide bonds and path length.

CD can be used to determine the secondary structure of a protein, as different structural compositions display characteristic spectra: α-helices give spectra with minima at 222 nm and 208 nm, and a maximum at 193 nm (Holzwarth and Doty, 1965); β-sheets give spectra with a minimum at 218 nm and a maximum at 195 nm (Greenfield and Fasman, 1969); disordered proteins, the content that is not α–helix or β–sheet/turn but can be just as rigid, give spectra with a minimum at around 195 nm and display low ellipticity above 210 nm (Venyaminov et al., 1993). Though unable to give structural detail on a residue by residue basis, unlike nuclear magnetic resonance (NMR) spectroscopy or X-ray crystallography for example, it does have the benefits of requiring little amounts of material and generally taking less time. Other requirements include the use of high-transparency quartz cuvettes, buffers should not contain anything optically active, and sample concentrations must be determined prior to

experimentation. Further information on CD protocols can be found in the Greenfield Nature Protocols review (Greenfield, 2006).

CD (Jasco J-815 spectrometer) data were collected by mixing the peptides in a 1:1 ratio in either the CD calcium buffer of HEPES (10 mM), Tween (0.005 %) and calcium chloride (2 mM), pH 7.0 or the CD EDTA buffer of HEPES (10 mM), Tween (0.005 %) and EDTA (2 mM), pH 7.0. First the peptides were investigated individually to determine the level of ordered structure, before mixing the peptides to observe any structural changes.

2.4.5 Ligation and thioester exchange

All ligation buffers thoroughly degassed whilst stirred (Buchi V-700 vacuum pump) and kept under argon, unless otherwise specified. For 1 ml of buffer a 5 ml round bottom flask was used. Any additives were introduced prior to setting the pH with 4 M NaOH or 12 M HCl as required. The final volume of ligation buffer was added first to the *N*-terminal peptide and then transferred to the *C*-terminal peptide. The reaction mixture was diluted in HPLC buffer A prior to injection onto analytical-HPLC, the specific dilutions detailed in individual sections.

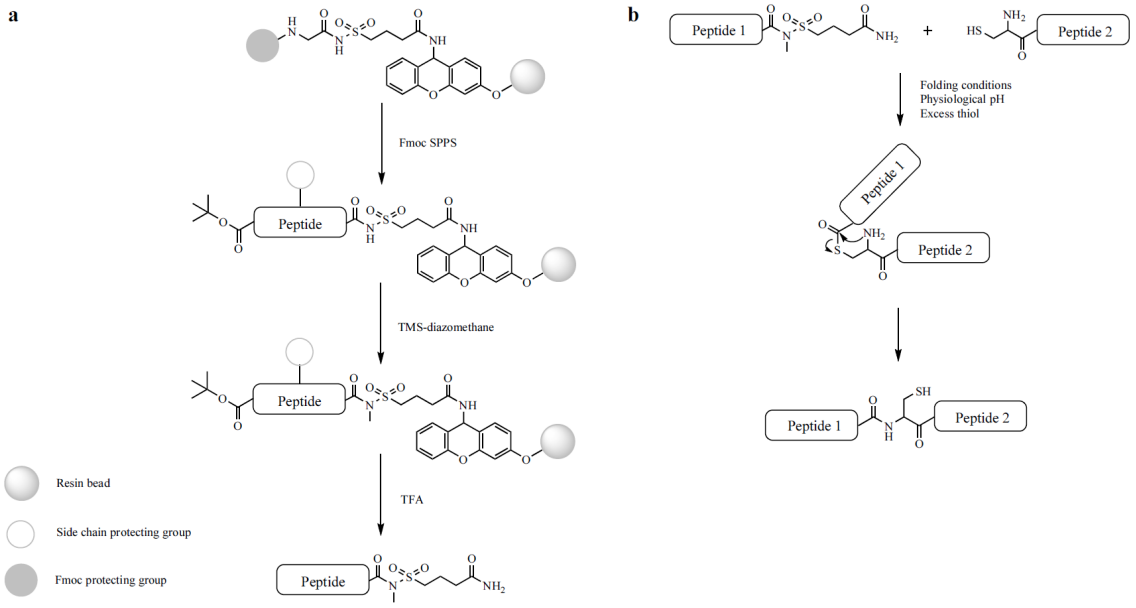
3. Native chemical ligation with peptide N-(Me)sulfonamides

3.1 Introduction

NCL is a very powerful technique for protein semisynthesis. Though peptide thioesters were traditionally synthesized by Boc-SPPS there has been extensive progress toward peptide thioester synthesis by the more prevalent Fmoc-SPPS (Mende and Seitz, 2011, Kang and Macmillan, 2010) for a number of reasons. The anhydrous hydrogen fluoride (HF) required for final deprotection and cleavage from the resin in Boc-SPPS is highly toxic, complex in handling and increasingly difficult to obtain. One problem in optimising an Fmoc-based method for synthesizing peptide thioesters is that the thioester linkage to the resin is sensitive to strong nucleophiles such as the piperidine used repeatedly throughout this method of synthesis. A number of approaches have been successfully developed, the most popular of which is probably the use of an acylsulfonamide safety-catch (acid and base stable) linker. Acylsulfonamides are resistant to alkaline hydrolysis because the base ionises the NH group which is acidic (Rudinger, 1963). Upon completion of peptide synthesis alkylation of a sulfonamide linker labilises it to nucleophilic cleavage, generating the fully-protected N-alkyl acylsulfonamide peptide (Kenner et al., 1971). The first application of this method for synthesizing peptide thioesters was developed by Pessi, where the thioester peptide was produced following activation and subsequent thiol displacement of the linker (Ingenito et al., 1999).

Published findings from our lab have shown the use of the sulfamylbutyryl linker to synthesise peptides for direct use in chemical ligation, the peptide N-

(Me)sulfonamide undergoing chemical ligation in the presence of excess thiols (Burlina et al., 2012). By using a double linker strategy of the sulfamylbutyryl linker in conjunction with a Sieber amide resin, a mildly activated group is achieved by methylation and cleavage without the need for the thioester conversion steps previously seen with this linker (Scheme 3.1). Trimethylsilyl-diazomethane (TMS-diazomethane) (Murphy et al., 2009) 2.0 M in hexanes was chosen for activation from previous work (Ingenito et al., 1999, Mezzato et al., 2005). This methylation reaction results in a mildly reactive peptide sulfonamide, reducing the risk of unwanted side reactions. During ligation reactions an intermediate thioester exchange step occurs between the peptide N-(Me)sulfonamide and 4-mercaptophenylacetic acid (MPAA) which is instantly captured by the *N*-terminal cysteine (Scheme 3.1).



Scheme 3.1: Preparation of a peptide N-(Me)sulfonamide for NCL

a. Double linker strategy for the preparation of a peptide N-(Me)sulfonamide. Following Fmoc SPPS treatment with TMS-diazomethane and TFA generates the activated, unprotected peptide N-(Me)sulfonamide.
b. In the presence of excess thiols the peptide N-(Me)sulfonamide undergoes thioester exchange before ligating to a peptide with an *N*-terminal Cys.

Adapted from Burlina et al, 2012.

Though not discussed in depth in this body of work our lab has used alkyl substitution of Cys to develop an additional method for the synthesis of thioester precursor peptides, synthesizing peptides with *C*-terminal α–methylcysteine (Burlina et al., 2014). Model ligations with LAPAGα-MeC and LAPAAα-MeC included a version with a LAPAGC control, before demonstrating ligation with BPTI, murine KC and a fragment of ColE1-Rop at Gly-Cys, His-Cys and Leu-Cys sites respectively.

3.2 Ligation with model peptides

This technique was first demonstrated using small model peptides in the form of AYRAA-N(Me)sufonamide and CFALRGWR-NH_2. Model peptides allow for replication of standard NCL conditions with their low molecular weight, use of denaturing buffer and high concentrations.

3.2.1 Preparation of peptides

AYRAA N-(Me)sulfonamide was synthesized on pre-loaded Fmoc-Ala-sulfamylbutyryl sieber amide resin (0.32 mmol g^{-1} loading, 0.1 mmol). Upon completion of the synthesis the resin was methylated by overnight treatment with TMS-diazomethane in hexanes (2.0 M). Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 0 – 30 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (26.8 mg, 32.4 % yield).

CFALRGWR was synthesized on PL-AMS resin with a rink amide linker (0.39 mmol g^{-1} loading, 0.1 mmol). Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 10 – 70 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried.

3.2.2 Ligation of AYRAA N-(Me)sulfonamide and CFALRGWR

AYRAA N-(Me)sulfonamide (0.65 mM, 0.36 mg) was added to CFALRGWR (1 mM, 0.42 mg) in degassed denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 7.5. The reaction was performed at 40 °C and monitored over time by analytical-HPLC: Chromolith RP-8 column with a gradient of 0 – 30 % B in 10 min and a flow rate of 3 ml min^{-1}. Ligation was complete in 2 h and the product AYRAACFALRGWR-NH_2 collected for MALDI-TOF analysis: observed mass = 1539.7 Da in positive reflector mode, calculated mass = 1539.8 Da.

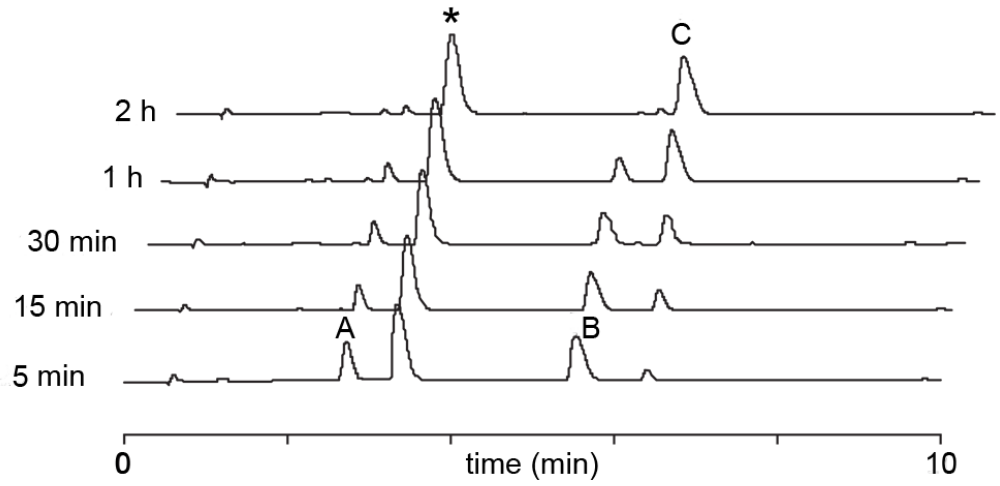


Figure 3.1: Ligation of AYRAA N-(Me)sulfonamide and CFALRGWR (Burlina, Morris et al, 2012)

Ligation conditions: peptides (10 mM), sodium phosphate (200 mM), GdnHCl (1 M), EDTA (2 mM), TCEP-HCl (50 mM), MPAA (60 mM), pH 7.5. HPLC conditions: Chromolith RP-8 column, 0 – 30 % B (TFA (0.1 %), AcN (90 %)) in 10 min, 3 ml min^{-1}. A = AYRAA N-(Me)sulfonamide, B = CFALRGWR, C = ligation product, * = MPAA.

3.3 Ligation with bovine pancreatic trypsin inhibitor

Having first demonstrated this form of ligation using model peptides, the technique was then demonstrated on a larger scale with peptides of higher molecular weight and with all side chain functionalities, in addition to lowered concentrations.

Bovine pancreatic trypsin inhibitor (BPTI) is a well-studied protein that has previously been totally chemically synthesized (Lu et al., 1998), where BPTI(1-37)αCOSR and BPTI(38-58) successfully ligated in 24 h. Here we have synthesized it as two fragments, BPTI(1-37) N-(Me)sulfonamide and BPTI(38-58). Using these peptides it should be possible to demonstrate the technique on a larger scale with peptides of higher molecular weight and with all side chain functionalities, in addition to lowered concentrations.

3.3.1 Preparation of peptides

BPTI(1-37) N-(Me)sulfonamide was synthesized on pre-loaded Fmoc-Gly-4-sulfamylbutyryl rink amide AM resin (0.26 mmol g^{-1} loading). Upon completion of the synthesis the resin was methylated by solvation in DCM and overnight treatment with TMS-diazomethane in hexanes (2.0 M). Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 35 % B in 5 min, 35 – 50 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (9 % yield).

BPTI(38-58) was synthesized on pre-loaded Fmoc-Ala-Wang resin (0.42 mmol g^{-1} loading). Cleavage from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 10 – 50 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (14 % yield).

3.3.2 Ligation of BPTI(1-37) N-(Me)sulfonamide and BPTI(38-58)

BPTI(1-37) N-(Me)sulfonamide (0.65 mM, 0.36 mg) was added to BPTI(38-58) (1 mM, 0.42 mg) in degassed denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 6.3. The reaction was performed at 40 °C and monitored over time by analytical-HPLC: Phenomenex RP-C18 column with a gradient of 0-70 % B in 25 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 3.2). Over time observe the formation of the ligation product BPTI, collected for characterization by MALDI-TOF: observed mass = 6519.4 Da in positive linear mode, calculated mass = 6517.6 Da.

This ligation was repeated on a larger scale and the product purified by preparative-HPLC. BPTI(1-37) N-(Me)sulfonamide (10 mg) was added to BPTI(38-58) (5.8 mg) in degassed denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM) additives, pH 6.3. The reaction was performed at 40 °C and after 10 h the product purified by preparative-HPLC: RP-C18 column with a gradient of 0 – 20 % B in 5 min, 20 – 65 % B in 30 min and a flow rate of 15 ml min^{-1} (2.4 mg, 15.2 % yield).

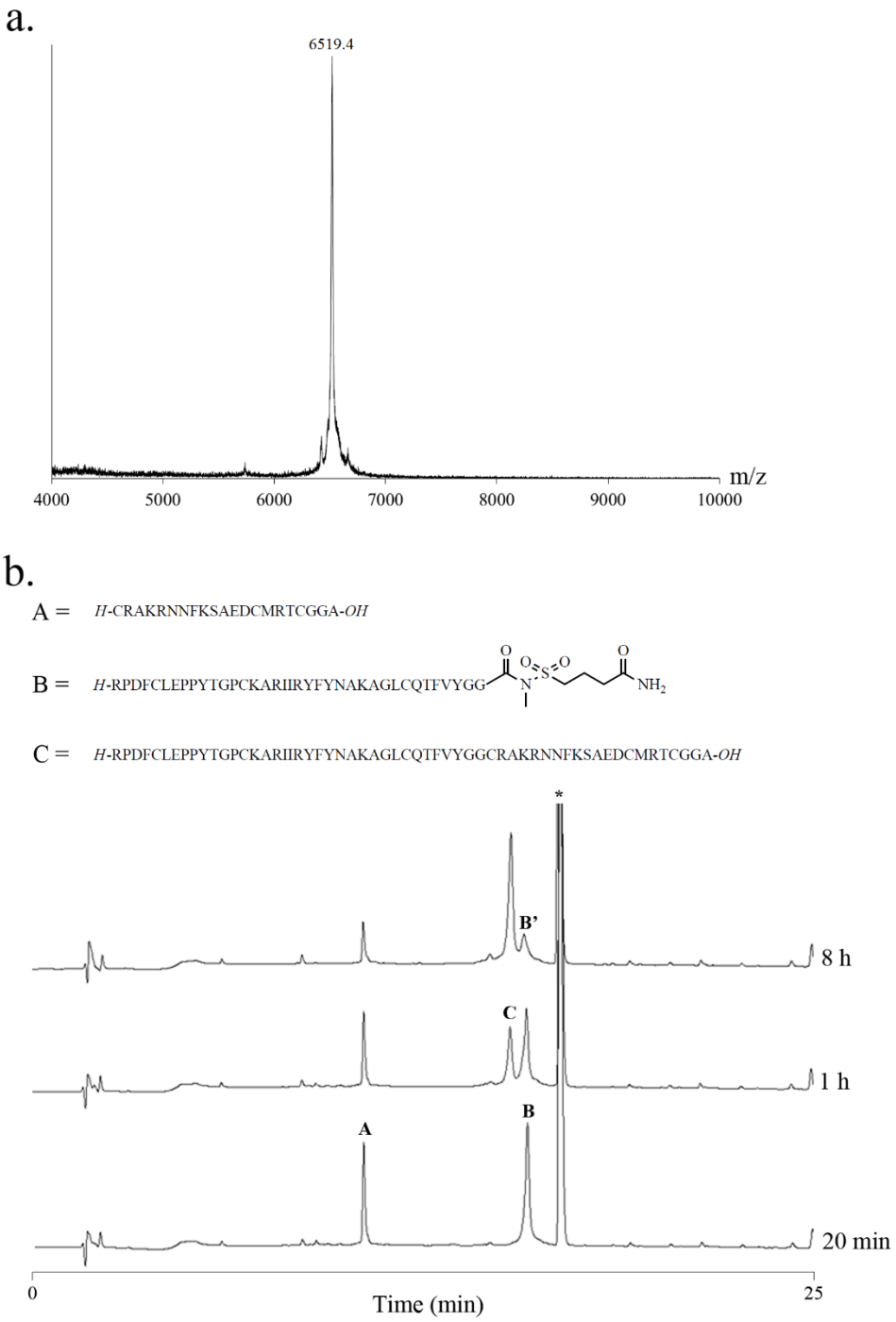


Figure 3.2: Ligation of BPTI(1-37) N-(Me)sulfonamide with BPTI(38-58)

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 6519.4 Da, calculated product mass = 6517.6 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), GdnHCl (6 M), sodium phosphate (200 mM), TCEP-HCl (50 mM) and MPAA (60 mM), pH 6.3, 40 °C. HPLC conditions: RP-C18 column, 0 – 70 % B (TFA (0.1 %), AcN (90 %)) in 25 min, 1 ml min^{-1}. A = BPTI(38-58), B = BPTI(1-37) N-(Me)sulfonamide, B' = BPTI(1-37), C = ligation product, * = MPAA (peak truncated).

3.4 Peptide N-(Me)sulfonamides ligate to *N*-terminal Cys peptides in the presence of excess thiols

This chapter overviews the work establishing a new and viable method for synthesizing peptide thioesters by Fmoc-SPPS. Peptide N-(Me)sulfonamides were synthesized using a double linker strategy to generate a mildly activated group following methylation and cleavage, without the requirement of a thioester conversion step previously undertaken with this linker. The double linker strategy allows for test cleavages during the synthesis of the peptide to monitor progress, including the final activation (Mezzato et al., 2005). Methylation resulted in a mildly reactive peptide sulfonamide, reducing the risk of unwanted side reactions. During ligation an intermediate thioester exchange step takes place between the peptide N-(Me)sulfonamide and the excess thiols present, before this thioester peptide then ligates to an *N*-terminal Cys peptide. This was initially carried out with model peptides, AYRAA N-(Me)sulfonamide with CFALRGWR (Figure 3.1), before BPTI was chosen as the target for a more thorough investigation. BPTI(1-37) N-(Me)sulfonamide ligated with clean analytical-HPLC traces to BPTI(38-58) at a Gly-Cys ligation site with MPAA as the thiol additive (Figure 3.2), these peptides based on the published total chemical synthesis. The BPTI(1-37) MPAA thioester is not observed during this ligation as it immediately undergoes transthioesterification with BPTI(38-58) to form the ligation product. The majority of the product had formed by 8 h, which makes the ligation with BPTI(1-37) N-(Me)sulfonamide described here comparable to the published total chemical synthesis.

The success of these peptide N-(Me)sulfonamides in ligation will now be applied to the split proteins discussed in later chapters, to establish their ability for use in the proposed labelling method.

3.5 Summary

A double linker strategy using the popular sulfamylbutyryl linker has afforded peptide N-(Me)sulfonamides that are able to directly undergo ligation. In the presence of excess thiols the peptide N-(Me)sulfonamide undergoes a thioester exchange step before ligating to *N*-terminal Cys-containing peptides, bypassing the previously separate thioester conversion step and allowing for monitoring throughout the synthesis. Ligation with model peptides paved the way for ligation with BPTI peptides, with results comparable to that of published NCL.

4. Conformationally-assisted ligation with RNase A

4.1 Introduction

Bovine RNase A was chosen as it was the classical system in which the first demonstrated example of reconstitution (Richards, 1958) and later conformationally-assisted ligation (Beligere and Dawson, 1999) were demonstrated. Additionally the S-peptide can be accessed by Fmoc-SPPS, and the larger S-protein is commercially available. Furthermore the activity of the authentic ligation product could be assayed using fluorescently labelled nucleotides, such as the RNaseAlert assay (Ambion).

RNase A, isolated from crystalline ribonuclease, was treated with subtilisin, cleaving the Ala-Ser bond 20 amino acids from the *N*-terminus of the protein (Richards, 1958). This liberated the 20 reside 'S-peptide' and the 104 residue 'S-protein', confirmed by amino acid analysis and Edman degradation (Richards and Vithayathil, 1959, Hirs et al., 1960), with 80-90 % conversion success. This complex, termed RNase S, was dissolved in trichloroacetic acid for separation of the components following centrifugation, the precipitant containing the S-protein and the supernatant the S-peptide. Activity was measured using either the acid-soluble method (Anfinsen et al., 1954) or the Kunitz spectrophotometric method (Kunitz, 1946). When the two components were mixed stoichiometrically 100 % of the original activity was observed. The reconstitution had no lag, and was immediate even for peptide concentrations under 1 μM. Later work established that the five *C*-terminal residues of S-peptide were not necessary for effective reconstitution, as stoichiometric amounts of this truncated S-peptide resulted in a return to full activity when mixed with S-protein as did the full-

length peptide (Potts et al., 1963). Work such as this on the optimisation of the system contributed to the decision of using RNase A in initial experiments.

Further key work that convinced us to use RNase A as a model system was the report of a conformationally-assisted ligation between an S-peptide thioester and the S-protein (Beligere and Dawson, 1999). Synthetic S-peptide thioester successfully ligated and formed an amide bond at the non-Cys *N*-terminus of S-protein under folding conditions with both peptides at 1 mM concentration, with the native Ser at the *N*-terminus of the S-protein. This report was novel as it demonstrated that *N*-terminal Cys wasn't necessary at the ligation site where the two fragments were held in close proximity. Ligation was complete in 10 h and the product possessed comparable activity to RNase A. Activity was assayed by cleavage of a fluorescent substrate (Kelemen et al., 1999, Song and Ismagilov, 2003).

4.2 Conformationally-assisted N-(Me)sulfonamide ligation and thioester exchange

We wanted to establish if we could use an S-peptide N-(Me)sulfonamide in place of a peptide thioester to ligate to S-protein with conformationally-assisted ligation. Because of the number of variants of S-peptide known and their different affinity to the S-protein we reasoned that this system would give us an idea of what affinity would be optimum for rapid ligation and eventually rapid labelling. We first wanted to repeat the work of Beligere and Dawson (Beligere and Dawson, 1999) and use conformational-assistance to ligate our synthetic S-peptide with the commercially available S-protein. Continuing from the previous chapter we decided to use peptide N-

(Me)sulfonamides instead of peptide thioesters for conformationally-assisted ligation with the split protein candidates.

4.2.1 Preparation of peptides

S-peptide N-(Me)sulfonamide

S-peptide N-(Me)sulfonamide was synthesized on pre-loaded Fmoc-Ala-sulfamylbutyryl sieber amide resin (0.32 mmol g^{-1} loading, 0.05 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Upon completion of synthesis the resin was methylated by solvation in DCM and overnight treatment with TMS-diazomethane in hexanes (2.0 M).

Integration of the methylated and unmethylated peaks by analytical-HPLC revealed 40 % conversion to methylated product (Figure 4.1 a.). Methylation of S-peptide showed no improvement after repeated treatments with methylation agent. One attempt to improve this was to include a Boc-protected amino acid at the *N*-terminus directly rather than independently coupling a Boc group on after the Fmoc-protected amino acid at this position was deprotected. If the Boc protection were incomplete then potentially the TMS-diazomethane would methylate at the *N*-terminus. Additionally we were concerned that some of the reagent used to introduce Boc may cleave the sulfonamide from the resin (Mezzato et al., 2005). Initially 5 eq Boc carbamate in minimal DMF was added to deprotected peptide at the end of the synthesis and left for 30 min with shaking, however this was later changed to *N*α,*N*ε-bis-Boc-L-Lys-*OH* for coupling at the *N*-terminus. Another modification that was introduced at the same time was the inclusion of a pseudoproline into the sequence. In addition to the possibility of multiple methylation sites, we reasoned that the solubility of the peptide and therefore

its accessibility to the methylating reagent could be a problem. Consequently peptide chain aggregation hinders the accessibility of the methylation agent to the peptide leading to a reduction in methylation efficiency. Introducing the pseudoproline Fmoc-Asp(OtBu)-Ser(ψMe,Mepro)-*OH* along with Nα,Nε-bis-Boc-L-Lys-*OH* into the sequence increased methylation to approximately 70 % (Figure 4.1 b.). Though the amount of methylated product improves with repeated overnight treatments, the question is whether this outweighs increased handling of the toxic methylation agent. An alternative methylation method in the form of methyl iodide and caesium carbonate was investigated, leaving the peptide overnight with 1 eq caesium carbonate and 2 eq methyl iodide in DMF (Hossan et al., 2012), but with poor results (data not shown). In an attempt to further improve the amount of methylated S-peptide using TMS-diazomethane the temperature of the methylation step was increased, first to 30 °C and then to 50 °C. These changes resulted in approximately 65 % and 66 % methylated product respectively (Figure 4.1 c.), which is no improvement. An overview of the effect of these modifications is given in Table 4.1

Following a synthesis modified in the above ways, the peptide was cleaved from the resin and purified by preparative-HPLC: RP-C18 with a gradient of 10 – 50 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (14 mg, 9.7 % yield).

Methylation attempt	Unmethylated (approximate %)	Methylated (approximate %)
Standard synthesis	60	40
Bis-boc-Lys/Pseudoproline	33	67
30 °C	35	65
50 °C	34	66

Table 4.1: Comparison of S-peptide sulfonamide methylation conditions

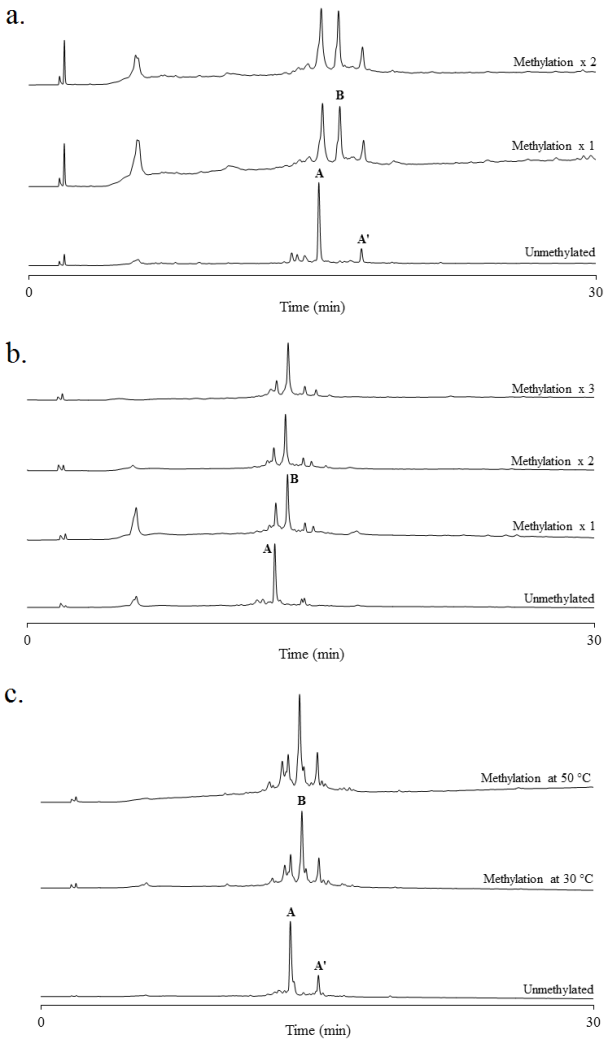


Figure 4.1: Preparation of S-peptide N-(Me)sulfonamide

a. Overnight treatment with TMS-diazomethane (2.0 M in hexanes) resulted in 40 % methylation, and this was not improved by subsequent overnight treatment. HPLC conditions: RP-C18 column, 10 – 50 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = S-peptide sulfonamide, A' = impurity, B = S-peptide N-(Me)sulfonamide.

b. Overnight treatment with TMS-diazomethane (2.0 M in hexanes) following the inclusion of Nα,Nε-bis-boc-L-Lys and Fmoc-Asp(OtBu)-Ser(ψMe,Mepro)-OH. Two subsequent overnight treatments improved the amount of methylated material. HPLC conditions: RP-C18 column, 0 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = S-peptide sulfonamide, B = S-peptide N-(Me)sulfonamide.

c. Overnight treatment with TMS-diazomethane (2.0 M in hexanes) at increased temperatures, at 30 °C and 50 °C. There appears to be no difference in the amount of product at elevated temperatures. HPLC conditions: RP-C18 column, 10 – 50 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = S-peptide sulfonamide, A' = impurity, B = S-peptide N-(Me)sulfonamide.

4.2.2 Ligation with thiophenol

The first attempt to repeat conformationally-assisted ligation used the non-denaturing buffer, in contrast to classical NCL which is performed in denaturing buffer containing GdnHCl, to exploit the secondary structure inherent to both components of RNase A, following the conditions reported by Beligere and Dawson (Beligere and Dawson, 1999).

S-peptide N-(Me)sulfonamide (10 mM, 0.4 mg) was added to S-protein (5 mM, 1 mg) in degassed non-denaturing buffer with thiophenol (2 % v/v), pH 7.5. The reaction was performed at RT and monitored by analytical-HPLC: RP-C18 column with a isocratic flow of 29 % B in 30 min and a flow rate of 1 ml min^{-1}.

Despite the successful identification of isocratic HPLC conditions to distinguish between S-protein and RNase A (Figure 4.2), ligation was not observed under these conditions (data not shown).

4.2.3 Ligation with 180 mM MPAA

MPAA was verified as a compatible thiol additive to the peptide N-(Me)sulfonamide approach in Chapter 3 and so the next attempt used this thiol additive in place of thiophenol as higher concentrations could be obtained to assist the ligation.

S-peptide N-(Me)sulfonamide (1 mM, 0.14 mg) was added to S-protein (1 mM, 0.68 mg) in degassed non-denaturing buffer with MPAA (180 mM), pH 7.5. A new gradient was developed for HPLC analysis, using the Chromolith RP8 column to establish a gradient over a quicker time frame and at a higher flow rate: Chromolith RP-

8 column with a linear gradient of 0 – 30 % B in 10 min and a flow rate of 3 ml min^{-1} (Figure 4.3).

There is hydrolysis of the S-peptide from the start of the reaction, though it does not appear to worsen with time. The product of the intermediate step, S-peptide MPAA thioester, is visible from as early as 30 min. The reconstituted ligation product, RNase A, is observed under these conditions and was collected and characterized by MALDI-TOF confirmation: observed mass = 13688.2 Da in positive linear mode, calculated mass = 13682.65 Da.

The experiment was repeated at higher MPAA concentration and lower pH, with no improvement of the ligation.

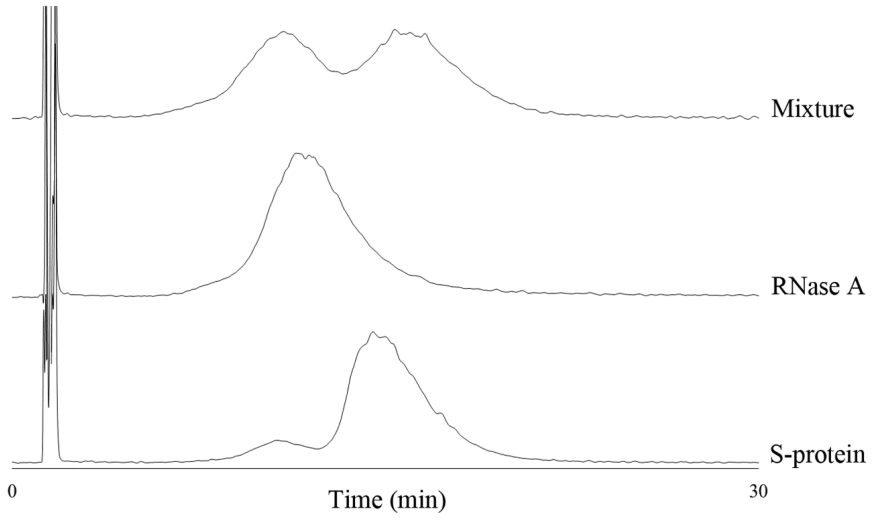


Figure 4.2: Development of an isocratic flow to distinguish between the S-protein and RNase A

An isocratic flow of 29 % B (TFA (0.1 %), AcN (90 %)) in 30 min with a flow rate of 1 ml min^{-1} on an RP-C18 column was identified, sufficiently establishing the retention time of the S-protein relative to the ligation product RNase A. A small additional peak in the S-protein trace corresponds to a nominal amount of intact RNase A.

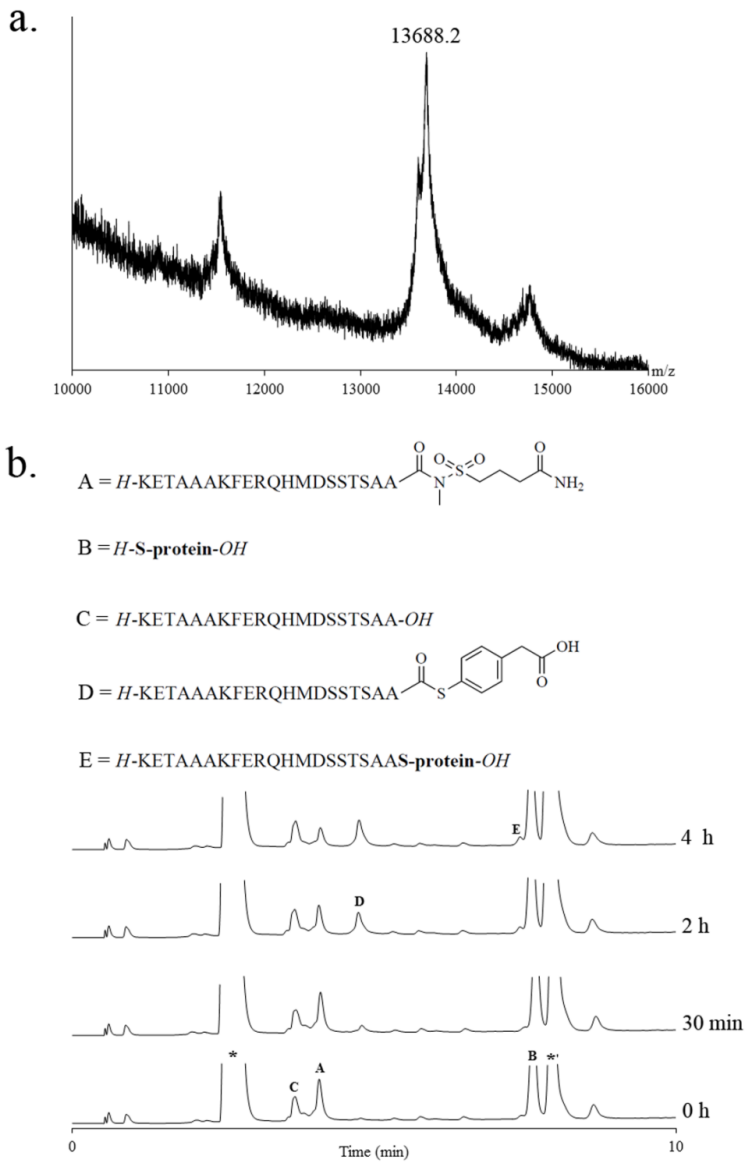


Figure 4.3: Conformationally-assisted ligation of S-peptide N-(Me)sulfonamide with S-protein, with 180 mM MPAA

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 13688.2 Da, calculated product mass = 13682.65 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), MPAA (180 mM), EDTA (2 mM), pH 7.5. HPLC conditions: Chromolith RP-8 column, 0 – 30 % B (TFA (0.1 %), AcN (90 %)) in 10 min, 3 ml min^{-1}. A = S-peptide N-(Me)sulfonamide, B = S-protein (peak truncated), C = S-peptide hydrolysis, D = S-peptide MPAA thioester, E = ligation product * = MPAA (peak truncated), *' = oxidised MPAA (peak truncated).

4.2.4 S-peptide N-(Me)sulfonamide exchange with MPAA

In all of the later ligation attempts we detected the peptide thioester intermediate, so it was decided to optimise the exchange conditions to raise the concentration of the more reactive species. Pre-exchanging the S-peptide N-(Me)sulfonamide for the MPAA thioester eliminates the requirement for thiol additives in the ligation step.

S-peptide N-(Me)sulfonamide (1 mM, 0.67 mg) was dissolved in degassed non-denaturing buffer with MPAA (300 mM), pH 8.0. The reaction was performed at RT and monitored by analytical-HPLC: Chromolith RP-8 column with a linear gradient of 0 – 30 % B in 10 min and a flow rate of 3 ml min^{-1} (Figure 4.4). The exchange product was collected and characterized by MALDI-TOF analysis: observed mass = 2317.3 Da in positive reflector mode, calculated mass = 2316.53 Da. After 5 h larger volumes were injected in order to collect the product, S-peptide MPAA thioester. Purified peptide freeze-dried (0.2 mg, 30 % yield).

A later attempt reduced the amount of thiol present, though this lengthened the exchange time. S-peptide N-(Me)sulfonamide (1 mM, 0.15 mg) was resuspended in degassed non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 0 – 50 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. Exchange is complete after 48 h.

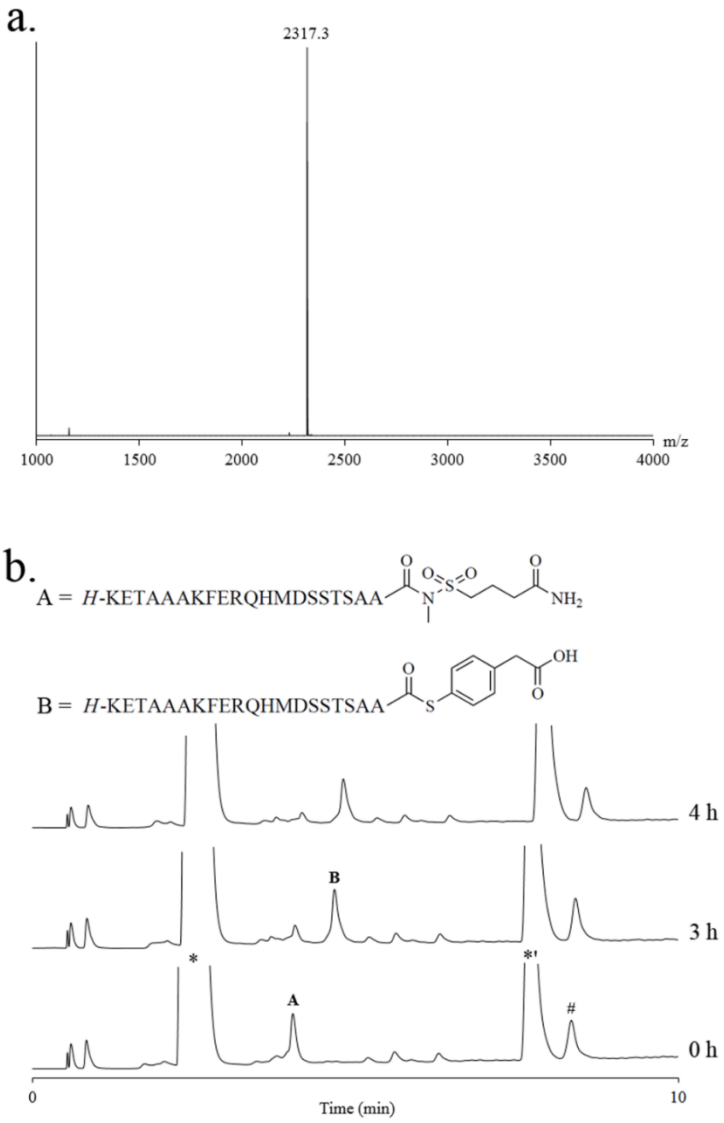


Figure 4.4: S-peptide N-(Me)sulfonamide exchange with MPAA

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 2317.3 Da, calculated product mass = 2316.53 Da.

b. Analytical-HPLC time course of exchange. Exchange conditions: peptide (1 mM), sodium phosphate (100 mM), MPAA (300 mM), pH 8.0. HPLC conditions: Chromolith RP-8 column, 0 – 30 % B (TFA (0.1 %), AcN (90 %)) in 10 min, 3 ml min^{-1}. A = S-peptide N-(Me)sulfonamide, B = S-peptide MPAA thioester, * = MPAA (peak truncated), *' = oxidised MPAA (peak truncated), # = MPAA impurity.

4.2.5 S-peptide N-(Me)sulfonamide exchange with MESNA

It has been reported that thiols with p*K*a values ≥ 6.5 make for better thioester exchange steps, and so the S-peptide N-(Me)sulfonamide was exchanged to the MESNA thioester (MESNA p*K*a 9.2) (Johnson and Kent, 2006).

S-peptide N-(Me)sulfonamide (1 mM, 0.17 mg) was dissolved in degassed non-denaturing buffer with TCEP-HCl (50 mM), MESNA (30 mM) and MPAA (60 mM), pH 7.0. The reaction was maintained at RT and monitored to completion by analytical-HPLC: RP-C18 column with a linear gradient of 0 – 50 % B in 30 min and a flow rate of 1 ml min^{-1} (Figure 4.5). Exchange is complete in 24 h, the product collected for MALDI-TOF confirmation: observed mass = 2291.0 in positive reflector mode, calculated mass = 2290.51 Da.

4.2.6 Ligation with pre-formed thioester

The ligation was now repeated with the exchanged S-peptide MPAA thioester.

S-peptide MPAA thioester (2 mM, 0.2 mg) was added to S-protein (1 mM, 0.5 mg) in degassed non-denaturing buffer, pH 7.5. The reaction was performed at RT and monitored by analytical-HPLC: Chromolith RP-8 column with a linear gradient of 0 – 35 % B in 10 min and a flow rate of 3 ml min^{-1} (Figure 4.6). Hydrolysis of the S-peptide MPAA thioester was detected at the t = 0 time point, but the ligation product was observed. The exact amount of ligation product is difficult to estimate due to the immediate hydrolysis of the S-peptide thioester, but there is relatively little. Again the product was confirmed by MALDI-TOF: observed mass = 13684.9 Da in positive linear mode, calculated mass = 13682.65 Da. An additional, unconfirmed peak was found to have an observed mass of 13599.3 Da in positive linear mode.

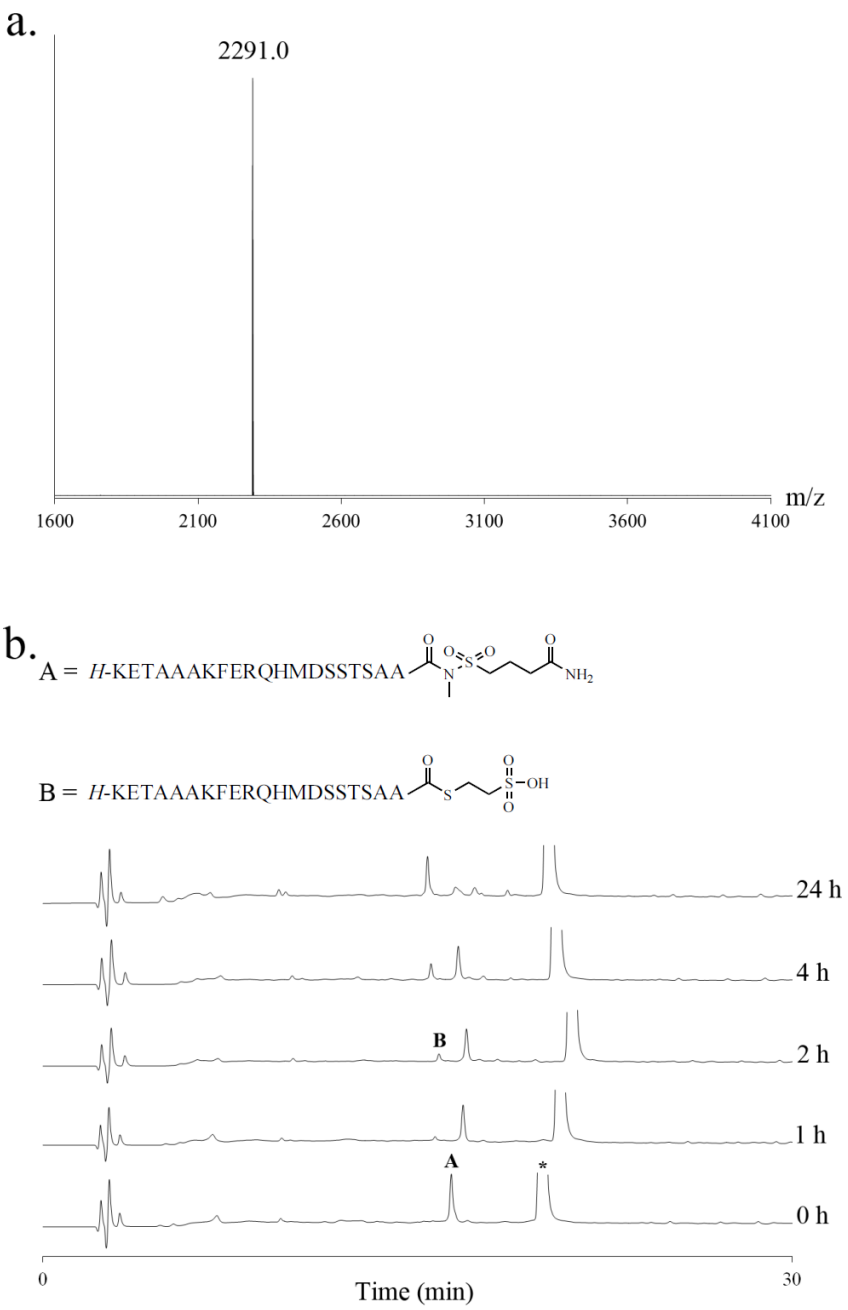


Figure 4.5: S-peptide N-(Me)sulfonamide exchange with MESNA

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 2291.0 Da, calculated product mass = 2290.51 Da.

b. Analytical-HPLC time course of exchange. Exchange conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MESNA (30 mM), MPAA (60 mM), pH 7.0. HPLC conditions: RP-C18 column, 0 – 50 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = S-peptide N-(Me)sulfonamide, B = S-peptide MESNA thioester, * = MPAA (peak truncated).

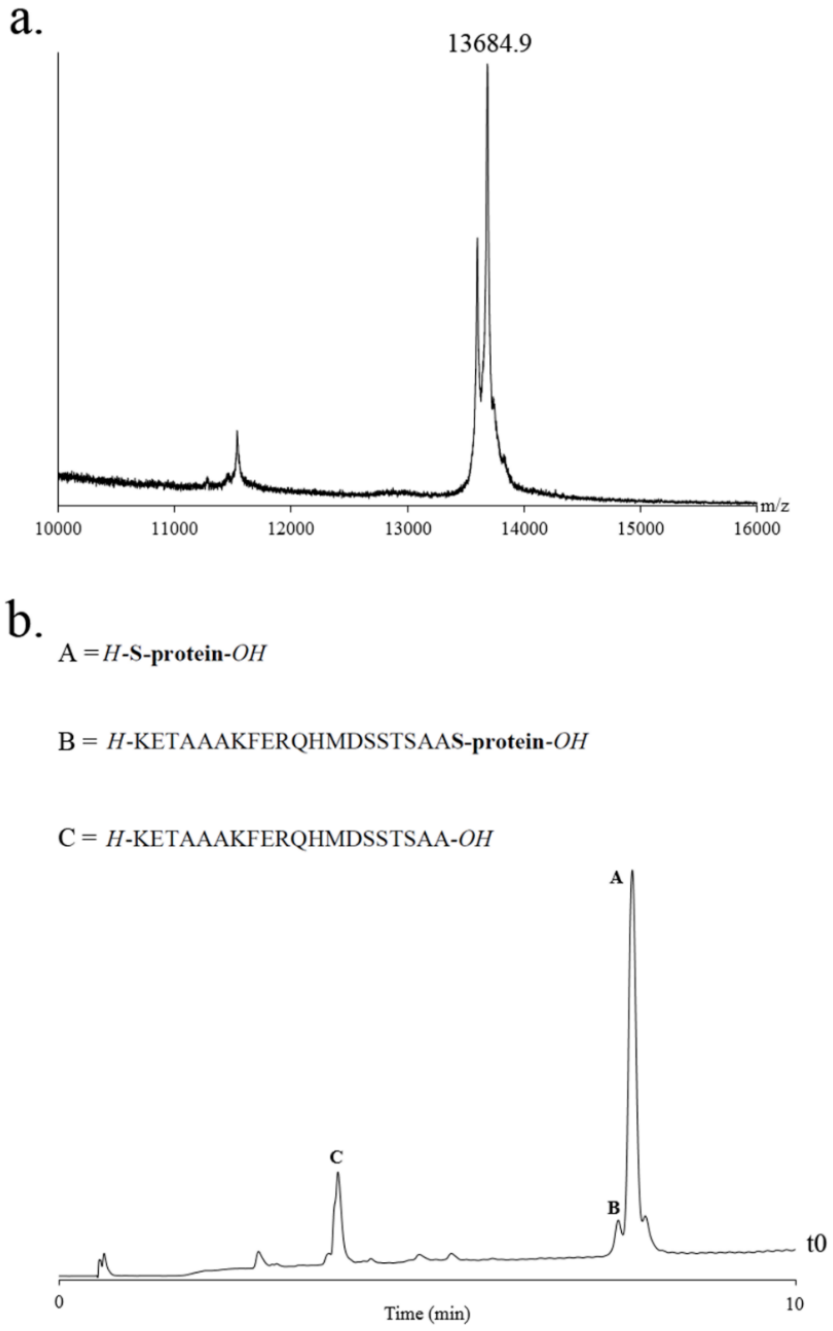


Figure 4.6: Conformationally-assisted ligation of pre-formed S-peptide MPAA thioester with S-protein

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 13684.9 Da, calculated product mass = 13682.65 Da.

b. Analytical-HPLC of ligation at t = 0. Ligation conditions: S-peptide MPAA thioester (2 mM), S-protein (1 mM), sodium phosphate (100 mM), pH 7.5. HPLC conditions: Chromolith RP-8 column, 0 – 35 % B (TFA (0.1 %), AcN (90 %)) in 10 min, 3 ml min^{-1}. A = S-protein, B = ligation product, C = S-peptide hydrolysis.

4.3 Biophysical characterization with BLI

With the results of the conformationally-assisted chemical ligation attempts between the RNase A partners, we examined different S-peptide variants from those reported to establish if any were more suitable for ligation than the full-length 20-mer investigated so far. Once the K_d of the full-length 20 residue S-peptide is established it can be compared to a number of variants: the 15-mer was chosen because of its smaller size; LB2 was reported in the literature for its high affinity for the S-protein, identified by phage display mutagenesis (Dwyer et al., 2001); we reasoned that a *C*-terminal Gly provided by an A20G mutation might assist ligation by minimizing steric hindrance. BLI was used to identify the affinity of these S-peptide variants for the S-protein, using the Octet system of FortéBio. Each S-peptide variant was biotinylated at its *N*-terminus for immobilization to streptavidin-coated sensors, and the interaction with the S-protein at its *C*-terminus was monitored. The peptide concentration was confirmed by UV/VIS spectrophotometry.

4.3.1 Preparation of peptides

Biotinylated S-peptide 20-mer

Biotinylated S-peptide 20-mer was synthesized on pre-loaded Fmoc-Ala-Wang resin (0.42 mmol g^{-1} loading, 0.1 mmol). Upon completion of synthesis biotin (0.2 M) was coupled at the *N*-terminus, confirmed with a negative Kaiser test. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 10 – 50 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (38.2 mg, 13.4 % yield).

Biotinylated S-peptide 15-mer

Biotinylated S-peptide 15-mer was synthesized on an AMS resin (0.1 mmol) with HMPAA linker. Upon completion of synthesis biotin (0.2 M) was coupled at the *N*-terminus, confirmed with a negative Kaiser test. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 10 – 50 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (14.4 mg, 5.92 % yield).

Biotinylated S-peptide LB2

Biotinylated S-peptide LB2 was synthesized on pre-loaded Fmoc-Ala-Wang resin (0.1 mmol). Upon completion of synthesis biotin (0.2 M) was coupled at the *N*-terminus, confirmed with a negative Kaiser test. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 20 – 50 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (2 mg, 5.85 % yield).

Biotinylated S-peptide A20G

Biotinylated S-peptide A20G was synthesized on pre-loaded Fmoc-Gly-Wang resin (0.2 mmol). Upon completion of synthesis biotin (0.2 M) was coupled at the *N*-terminus, confirmed with a negative Kaiser test. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 10 – 50 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (7 mg, 4.94 % yield).

4.3.2 S-peptide 20-mer with S-protein

Biotinylated S-peptide 20-mer (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *N*-terminus and is introduced to varying concentrations of S-protein (0.9 – 15.7 μM), both peptides in a PBS buffer with TCEP-HCl (0.5 mM), pH 7.4. A K_d value of 730 nM ± 130 nM was observed (Figure 4.7).

4.3.3 S-peptide 15-mer with S-protein

Biotinylated S-peptide 15-mer (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *N*-terminus and is introduced to varying concentrations of S-protein (0.9 – 15.7 μM), both peptides in a PBS buffer with TCEP-HCl (0.5 mM), pH 7.4. A K_d value of 1.65 μM ± 0.16 μM was observed (Figure 4.7).

4.3.4 S-peptide LB2 with S-protein

Biotinylated S-peptide LB2 (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *N*-terminus and is introduced to varying concentrations of S-protein (0.65 – 11.9 μM), both peptides in a PBS buffer with TCEP-HCl (0.5 mM), pH 7.4. A K_d value of 450 nM ± 170 nM was observed (Figure 4.8).

4.3.5 S-peptide A20G with S-protein

Biotinylated S-peptide A20G (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *N*-terminus and is introduced to varying concentrations of S-protein (0.65 – 11.9

μM), both peptides in a PBS buffer with TCEP-HCl (0.5 mM), pH 7.4. A K_d value of 900 nM ± 210 nM was observed (Figure 4.8).

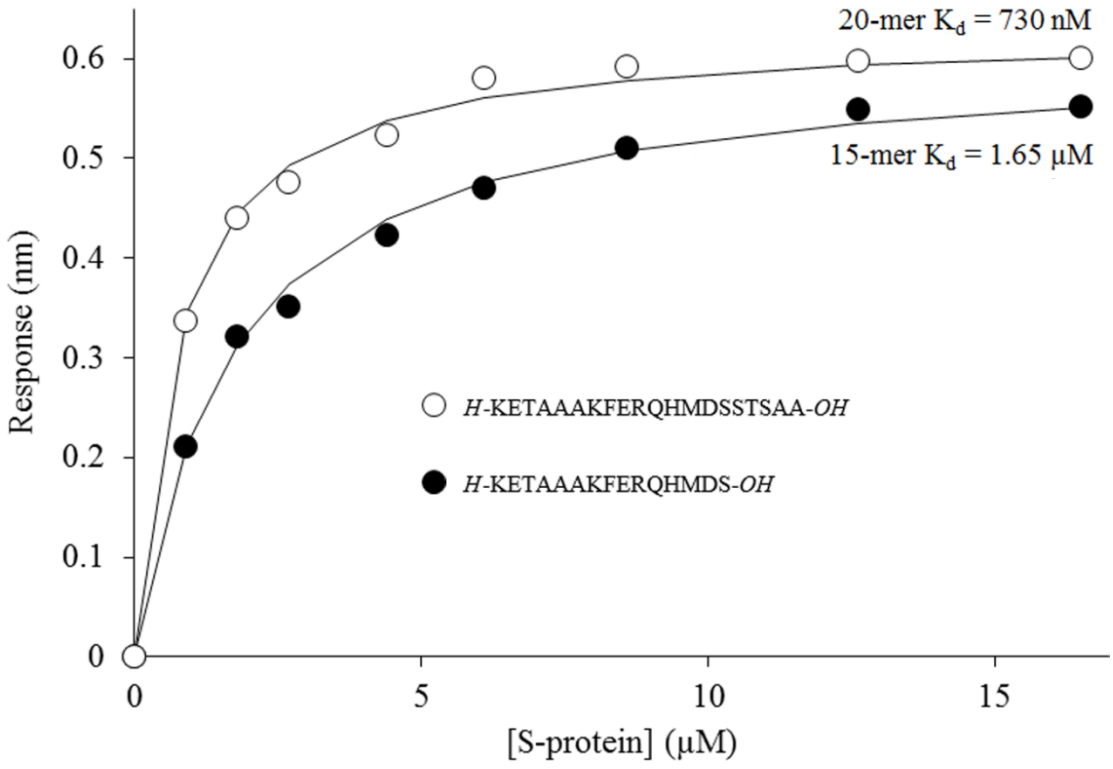


Figure 4.7: BLI comparison of 20-mer and 15-mer S-peptides

The full-length 20-mer S-peptide is determined to have a K_d value of 730 nM ± 130 nM, white circles, in comparison to a K_d value of 1.65 μM ± 0.16 μM for the 15-mer S-peptide, black circles. In both cases the [S-protein] is plotted against the response detected.

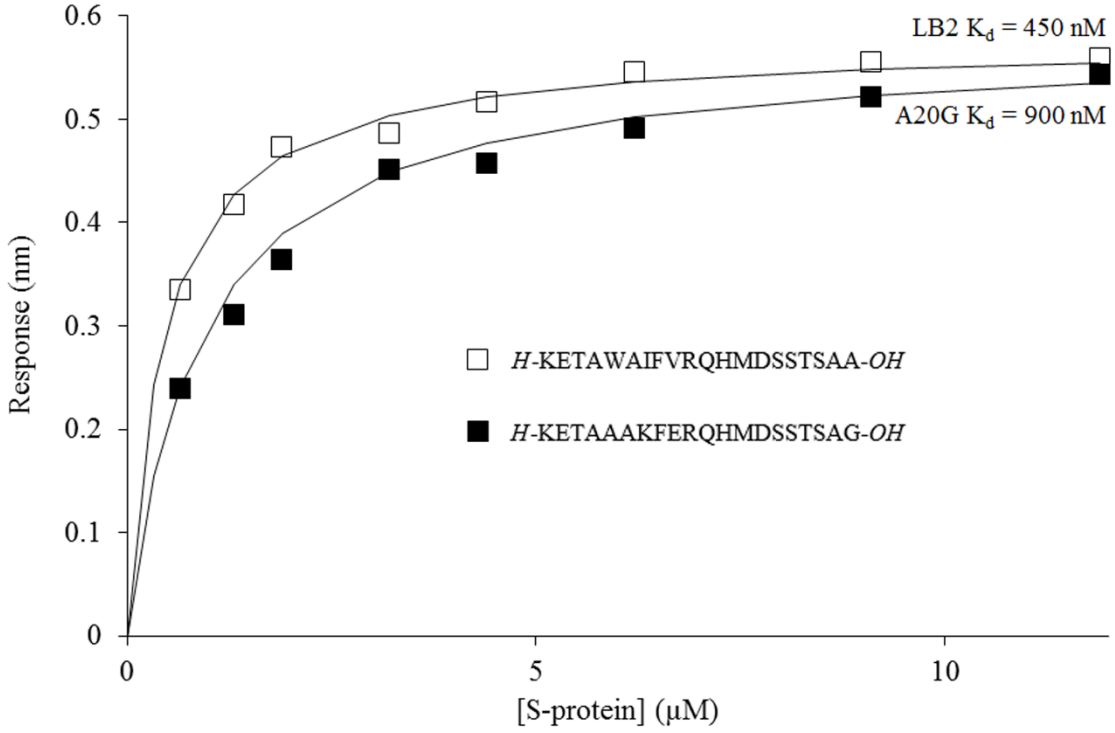


Figure 4.8: BLI comparison of LB2 and A20G S-peptides

The LB2 S-peptide is determined to have a K_d value of 450 nM ± 170 nM, white squares, in comparison to a K_d value of 900 nM ± 210 nM for the A20G S-peptide, black squares. In both cases the [S-protein] is plotted against the response detected.

S-peptide variant	K_d
20-mer	730 nM
15-mer	1.65 µM
LB2	450 nM
A20G	900 nM

Table 4.2: K_d values for S-peptide variants obtained by BLI

4.4 Attempting conformationally-assisted reactions with RNase A

4.4.1 Conformationally-assisted ligation

RNase A is the classic system for the study of protein folding and reconstitution, and we therefore chose it for our initial ligation studies. The first experiments with RNase A revolved around the conformationally-assisted ligation of the synthetic S-peptide N-(Me)sulfonamide with commercial S-protein. Conformationally-assisted thioester ligation has previously been reported with RNase A by Beligere and Dawson and if we can replicate this, we can show the suitability of this reaction for the proposed labelling method. We began by attempting to replicate selected experiments using a peptide N-(Me)sulfonamide for ligation as established as a thioester substitute in Chapter 3. Though chosen for initial investigation the S-protein, at 104 residues in length, would not make this an ideal system for protein labelling as proposed as it is too large a tag with multiple disulfide bonds. Additionally ligation would result in a functional ribonuclease, which would not necessarily be compatible with cellular labelling.

The single overnight treatment with the methylation agent TMS-diazomethane that was sufficient with the previous model peptides in Chapter 3 now only yielded approximately 40 % methylated product. Following a synthesis with both the *N*α,*N*ε-bis-Boc-L-Lys-*OH* and Fmoc-Asp(OtBu)-Ser(ψMe,Mepro)-*OH* the amount of methylated product increased to around 67 %. With this improvement it was wondered whether increasing the temperature might help as temperature can often decrease aggregation. Conducted in a fume hood, as must always be done (Murphy et al., 2009), overnight treatment with TMS-diazomethane was carried out at first 30 °C and then 50 °C. At 30 °C ~ 65 % was methylated, and ~ 66 % at 50 °C. As illustrated in Table 4.1

changing the temperature of the reaction had no significant effect on methylation, whereas the initial introduction of the pseudoproline increased the amount of methylated product by 25 % (Figure 4.1). The addition of the pseudoproline to the synthesis has improved on-resin solubility of the peptide and improved methylation, and subsequent temperature increases has not had any further effect. To improve solubility another method would be to add N-(2-hydroxy-4-methoxybenzyl) (Hmb)-protected amino acids into the synthesis, ideally every six or so residues (Johnson et al., 1993, Hyde et al., 1994, Simmonds, 1996), preceding the problematic area to disrupt the problematic hydrogen-bonding. Like pseudoprolines, these are easily cleaved at the end of the synthesis during the TFA treatment.

The first ligation attempt with S-peptide N-(Me)sulfonamide was conducted under folding conditions with thiophenol additive to replicate the conditions published for conformationally-assisted ligation with these peptides (Beligere and Dawson, 1999). An isocratic flow was established to distinguish between the ligation product and RNase A, predicting a small shift in the HPLC trace when ligating a small peptide onto a much larger one. 29 % B in 30 min on the RP-C18 column with a flow rate of 1 ml min^{-1} was found to sufficiently separate the two (Figure 4.2). When no ligation product was observed under these conditions it was decided to change the thiol from thiophenol to the more soluble MPAA to emulate the ligation conditions in Chapter 3, as this increased the concentration of active thioester in aqueous solution (Johnson and Kent, 2006), however we were concerned it would denature S-protein by reducing disulfides. Though the isocratic flow distinguishes between the S-protein and the product the analytical-HPLC used did not have a column oven, and fluctuating temperature resulted in shifting of the peaks complicating identification leading to the identification of a suitable gradient using the Chromolith column. Here we see the first indication of

ligation product, with a small peak corresponding to RNase A adjacent to the S-protein peak (Figure 4.3). We also observe the thioester intermediate, where S-peptide N-(Me)sulfonamide has exchanged to S-peptide MPAA thioester. Hydrolysis of S-peptide N-(Me)sulfonamide is visible from the t = 0 trace, along with oxidation of MPAA. Increasing the amount of thiol and lowering the pH had no effect.

In each ligation attempt the intermediate peptide thioester is visible by analytical-HPLC. Exchange was carried out separately to ligation, as the use of the pre-formed S-peptide MPAA thioester should increase the rate by avoiding this rate-limiting step. Aryl thiols such as MPAA are good leaving groups for the final ligation step because of the stability of the thiophenolate anion, and pre-formed aryl thioesters give quicker ligations (Johnson and Kent, 2006). The first exchange used 300 mM MPAA and was complete in 5 h (Figure 4.4). In a later exchange MPAA was lowered to 60 mM and TCEP-HCl was added to the exchange buffer in an attempt to minimise hydrolysis of the S-peptide N-(Me)sulfonamide and oxidation of MPAA as no S-protein was present. Exchange was not complete until 48 h and this lengthened time allowed for competing hydrolysis of the starting material, with the hydrolysis product becoming more substantial from 24 h. A third exchange was conducted with MESNA as an alternative thiol with the reasoning that hydrolysis with a less activated thioester would be reduced, and with recent literature demonstrating its use in peptide ligation (Adams et al., 2013). This yielded the S-peptide MESNA thioester in 24 h with less hydrolysis (Figure 4.5). The product of the first exchange reaction was purified for ligation to the S-protein, hoping that the use of a pre-exchanged peptide thioester would not only increase the speed of ligation but that obviating the presence of thiols that could reduce disulfide bonds in the S-protein could also improve efficiency. Additionally we wondered if the sulfonamide component might be sterically hindering ligation. Under

such conditions the ligation product RNase A is observed but there is no real improvement on the amount of product generated without the pre-formed thioester (Figure 4.6).

Unlike the ligations in Chapter 3, the ligations with RNase A cannot use TCEP-HCl. This is because the S-protein contains four disulfide bridges (Anfinsen et al., 1954, Sela et al., 1957) that need to be maintained for its secondary structure, and the addition of TCEP-HCl would reduce them and denature the protein. One of the main problems encountered was the shift of the large MPAA peak due to its oxidation, along with issues of hydrolysis with the S-peptide N-(Me)sulfonamide and thioester peptides. Another problem common to the attempts noted here is that the product forms such a small peak adjacent to the large S-protein peak by analytical-HPLC, and is difficult to separate by varying the gradient used. It is worth noting that in order to confirm the identity of the suspected product peak it should be isolated and assayed for activity, to confirm that S-peptide has not ligated to an S-protein side chain. None of the conditions examined yielded enough product to do this, and losses from hydrolysis made it impractical to continue. The reported conformationally-assisted ligation was limited to a single table entry with no further conditions give in the supplementary materials (Beligere and Dawson, 1999) and we concluded that their results was probably based on misinterpretation of mass spectrometry. Without experimental details for the published conformationally-assisted ligation with RNase A many different conditions have been attempted with limited success. Considering RNase A was not an ideal split protein system for labelling because of size and disulfide bridges, we decided to investigate other more promising, smaller split proteins with greater reported affinity.

Three different thiols were used in these ligation attempts, as the choice of thiol is important. A study to explore thiol exchange in peptide-α-thioesters using model

thioester peptides treated with an excess of a competing thiol showed that the use of a thiol additive that gives a more reactive leaving group can increase the rate of ligation through *in situ* transthioesterification (Dawson et al., 1997). Ligations with 4 % thiophenol were found to be complete in 7 h, whereas ligations with 4 % benzyl mercaptan were only 25 % complete in this time. Benzyl mercaptan (pK_a 9.4) is a strong nucleophile but alkyl thioesters are less reactive, whereas thiophenol (pK_a 6) is a poor nucleophile that provides highly activated thioesters. The more activated the ester the more susceptible it is to potential hydrolysis, as witnessed with the work in this chapter. Pre-formed thioesters ligate quicker by eliminating the thioester exchange step, and different thiols are better for the two NCL steps of exchange and transthioesterification. One of the most popular thiol additives in NCL is MPAA, having no offensive odour and higher solubility compared to many other thiol additives (Johnson and Kent, 2006). Aryl thiols, such as MPAA (pK_a 6.6), provide thioesters that are very good leaving groups during the transthioesterification step of a ligation, though exchange can be rate-limiting. Alkyl thiols, such as MESNA (pK_a 9.2), provide alkyl thioesters that make for poor leaving groups due to their low reactivity, though the thioester exchange step with such a thiol is not rate-limiting in the ligation. The poor basicity of the alkyl thiolate makes it a poor leaving group in contrast to the conjucated thiophenolate anion. As such the addition of MESNA to a ligation with MPAA slows the reaction by competing during the exchange step. Model ligation with MPAA and turkey ovomucoid third domain (OMTKY3) established MPAA as one of the more popular thiol choices compared to other thiols, for the reasons described (Johnson and Kent, 2006).

4.4.2 Biophysical characterization with BLI

Under all conditions tested within this chapter the hydrolysis and oxidation were competing side reactions that caused problems. Having varied the pH, thiol, concentration of thiol and starting material, and the use of a pre-formed thioester peptide, we next thought to optimise the S-peptide. First binding of the full-length 20-mer S-peptide to S-protein has been used, and the first step was to establish the affinity of it for the S-protein for comparison to that reported in the literature. BLI was the chosen technique for studying this interaction, after surface plasmon resonance (SPR) results in the literature established a Kd of 599.0 nM ± 68 nM (Dwyer et al., 2001). Though SPR had been used in this paper we have access to a FortéBio Octet red system within the division, allowing for a straightforward collaboration. BLI allows for calculation of K_d in a similar manner to SPR, the main difference being that BLI uses shaking rather than the fluidics of SPR. This same piece of literature also details the high-binding LB2 variant with an observed K_d of 5.4 nM ± 1.3 nM (Dwyer et al., 2001), and so this variant was synthesized for analysis by BLI. The truncated 15-mer S-peptide noted in the literature for its comparable affinity for the S-protein as the wt (Potts et al., 1963) was also synthesized, along with a variant with the mutation A20G. The latter was synthesized in case the *C*-terminal Ala used in the ligation attempts provided too much steric hindrance. In each instance biotin was coupled to the *N*-terminus of the peptide in order for the S-peptide to be immobilised during experimentation, binding streptavidin-coated sensors to monitor the interaction at its *C*-terminus. By testing three alternative sequences we can establish if there are better S-peptides than that currently used for ligation. See Table 4.2 for an overview of our BLI results.

The K_d of the interaction between the full-length 20-mer S-peptide was determined as 730 nM (Figure 4.7), which is in agreement with the published value of

600 nM (Dwyer et al., 2001). The truncated complex with the 15-mer S-peptide was found to have a K_d of 1.65 μM (Figure 4.7), a reduced affinity compared to the original S-peptide. With this variant ruled out, the high-binding LB2 peptide provided a K_d of 450 nM (Figure 4.8), suggesting an improved interaction to the 20-mer. The literature reports that the LB2 variant binds 110-fold more tightly than the wt with a K_d of 5.4 nM, but we were unable to observe an affinity similar to that. We wondered if this difference could be due to the immobilised component in both experiments. In the published SPR work the S-protein is immobilised whereas we immobilise the S-peptide, and we wondered if as the S-protein is the larger component whether the interaction was obstructed by its proximity to the sensor. Perhaps the immobilized S-peptide is too close to the surface of the sensor and is less accessible to the S-protein, relayed to us as reduced affinity, which could be tested for by the addition of a linker region between the S-peptide and the biotin at its *N*-terminus. Finally the S-peptide with a *C*-terminal Gly was identified as having a similar affinity for the S-protein as the full-length peptide, with a K_d of 900 nM (Figure 4.8). Based on these experiments the only variant that could potentially improve ligation attempts with the S-protein is the LB2 peptide, and the possibility of synthesizing the S-peptide LB2 N-(Me)sulfonamide for further conformationally-assisted ligation work has been left as a potential avenue of future work. It was decided to change the split protein system for a smaller one with a higher affinity.

4.5 Summary

Bovine RNase A was chosen as the first split protein system for investigation. Conformationally-assisted ligation was observed between the S-peptide N-(Me)sulfonamide and the S-protein, but hydrolysis of various S-peptide thioesters was an issue and ligation itself was limited. A number of thiol additives were explored, along with the use of a pre-formed thioester peptide. With limited success alternative split proteins were studied.

5. Conformationally-assisted ligation with CI2

5.1 Introduction

With only limited conformationally-assisted ligation of RNase A components, CI2 was chosen as the next candidate. If this system proved successful for translation to the proposed labelling method it would be more suitable than RNase A: firstly it has a much smaller *C*-terminal peptide, the portion to be expressed on the target peptide, and secondly reconstitution would not generate a functional ribonuclease. As discussed in Chapter 1 there is an appreciable amount of literature available, and like RNase A is a classic protein for folding studies.

The key work upon which the following experiments were based is that of Beligere and Dawson (Beligere and Dawson, 1999, Beligere and Dawson, 2000), where fragments of CI2 were employed to achieve regioselective coupling based upon their conformation. This relates to previous work on RNase A via the observation that both fragments of the protein maintain folding after subtilisin digestion (Richards and Vithayathil, 1959). They hypothesized that the conformation of CI2(1-39) and CI2(40-64) would bring their interaction site *C*- and *N*-termini respectively into close proximity and induce ligation. In order to avoid ligation with a β–branched amino acid CI2(1-39)T39D was synthesized, a substitution shown to have little impact on folding and stability (Itzhaki et al., 1995). After demonstrating ligation between CI2(1-39)T39D and CI2(40-64)M40C under denaturing conditions, they were influenced by such reactions as the ligation between coiled-coil peptides (Lee et al., 1996) to consider the effect of conformation on rate. Utilising the peptide conformation and the proximity of the two termini at the ligation site the rate was significantly improved under folding conditions not only with the native Thr but also with the native Met and no Cys, a ligation which

was not observed to occur under denaturing conditions. Later work also assembled a 4-ester CI2 analogue by conformationally-assisted chemical ligation where four residues in the α–helix of CI2(1-39) were substituted with α–hydroxy acids, disrupting three neighbouring hydrogen bonds (Beligere and Dawson, 2000). This, along with the T39D substitution, would suggest some flexibility in choice of amino acids: in order to ligate these peptides must be folded, therefore these substitutions are permitted.

5.2 Conformationally-assisted ligation of CI2(1-39)T39G N-(Me)sulfonamide with CI2(40-64)

The work described by Dawson and colleagues demonstrated the influence of peptide conformation. Initial work with CI2 involved repeating the work described above, but using CI2(1-39)T39G N-(Me)sulfonamide. It was hypothesized that as substituting the β–branched Thr for Asp had no negative impact on ligation, substituting to Gly for minimal steric hindrance would cause no problems. CI2(40-64) was synthesized with the native Met at its *N*-terminus.

5.2.1 Preparation of peptides

CI2(1-39)T39G N-(Me)sulfonamide

CI2(1-39)T39G N-(Me)sulfonamide was synthesized on PL rink resin (0.39 mmol g^{-1} loading, 0.1 mmol) with Fmoc-Gly-4-sulfamylbutyric acid linker (3 eq) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Upon completion of synthesis the resin was methylated by solvation in DCM and overnight treatment with TMS-diazomethane in hexanes (2.0 M). Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 30 %

B in 5 min, 30 – 60 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (16 mg, 3.12 % yield).

CI2(40-64)

CI2(40-64) was synthesized on pre-loaded Fmoc-Gly-Wang resin (0.76 mmol g^{-1} loading, 0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried.

5.2.2 Ligation with MPAA

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.47 mg) was mixed with CI2(40-64) (1 mM, 0.32 mg) in non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 6.3. The reaction monitored by analytical-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. Over time observe the formation of CI2(1-39)T39G MPAA thioester, but no ligation product (Figure 5.1). The thioester intermediate was collected for confirmation by MALDI-TOF: observed mass = 4427.3 Da in positive reflector mode, calculated mass = 4424.43 Da.

Another attempt replicated these conditions but the pH was raised from 6.3 to 7.5. As with the initial attempt, we observe the formation of CI2(1-39)T39G MPAA thioester but no ligation product (data not shown).

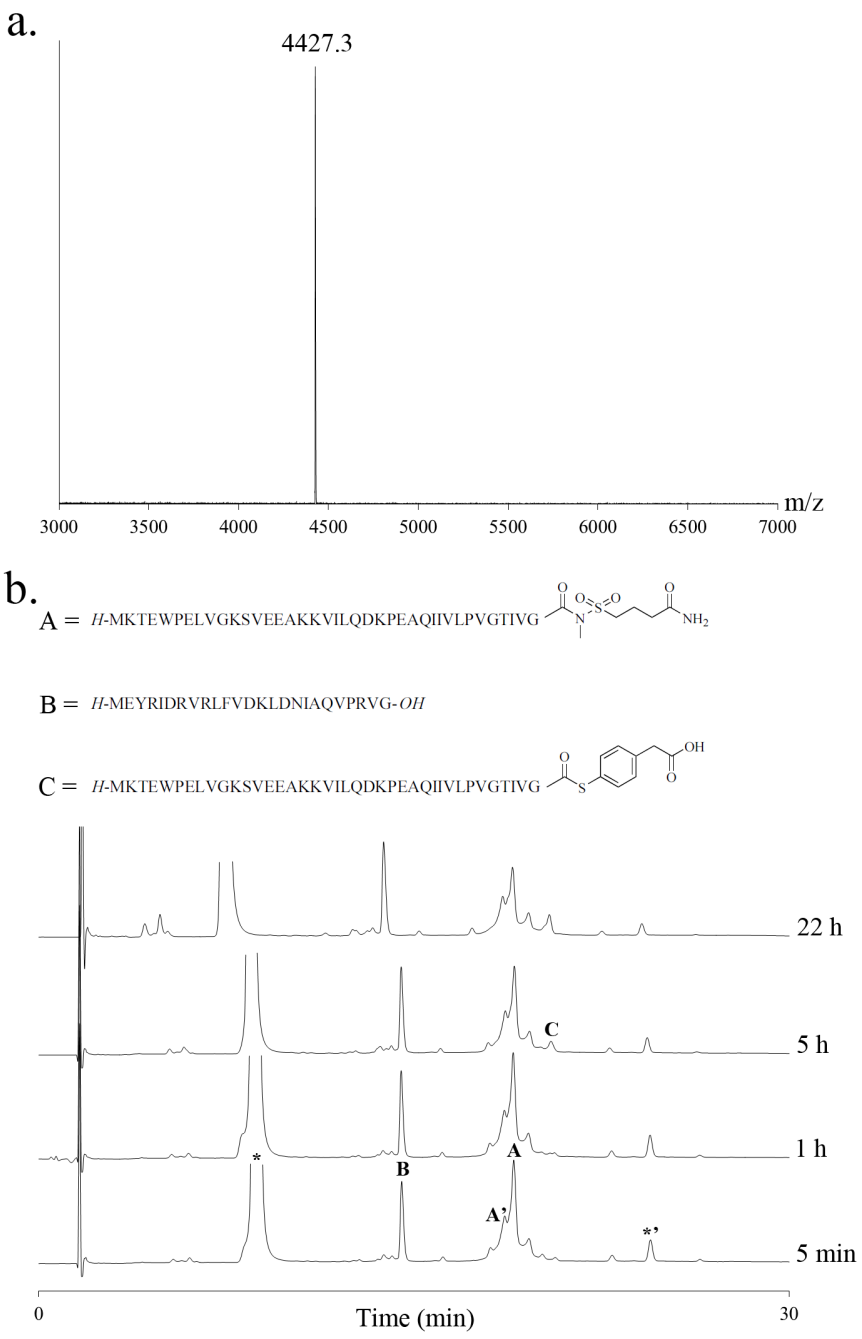


Figure 5.1: Conformationally-assisted ligation of CI2(1-39)T39G N-(Me)sulfonamide with CI2(40-64), with MPAA

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 4427.3 Da, calculated product mass = 4424.43 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (60 mM), pH 6.3. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(1-39)T39G N-(Me)sulfonamide, A' = CI2(1-39)T39G impurity, B = CI2(40-64), C = CI2(1-39)T39G MPAA thioester, * = MPAA (peak truncated), *' = MPAA impurity.

5.2.3 Thioester exchange with MPAA and benzyl mercaptan

Ligation attempts so far have shown thioester exchange with CI2(1-39)T39G N-(Me)sulfonamide, and it was decided to take advantage of this. By pre-exchanging we can introduce the CI2(1-39)T39G thioester directly to CI2(40-64) under the published folding conditions, without the requirement for thiol addition and bypassing the thioester exchange step to try to increase rate. Firstly the conditions of the above ligation attempts was used, with MPAA as the sole thiol.

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.24 mg) was resuspended in non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 7.0. The reaction was monitored by analytical-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. Exchange is taking place, though perhaps slower than ideal.

It is decided to add benzyl mercaptan (10 % v/v) before leaving overnight, being a strong nucleophile this should make the less reactive CI2(1-39)T39G benzyl mercaptan thioester which would be more stable if lengthy exchange times are to be expected. Following this we see in the 19 h trace that most of the starting material has been used and has converted to the benzyl mercaptan thioester, collected and characterized by MALDI-TOF analysis: observed mass = 4383.4 Da in positive reflector mode, calculated mass = 4380.22 Da (Figure 5.2).

Having converted CI2(1-39)T39G N-(Me)sulfonamide to the CI2(1-39)T39G benzyl mercaptan thioester, the exchange reaction was repeated with benzyl mercaptan alone. Though there is exchange product in the overnight time point it appears to be progressing slower than the previous reaction, suggesting that MPAA must be present as a catalyst (data not shown).

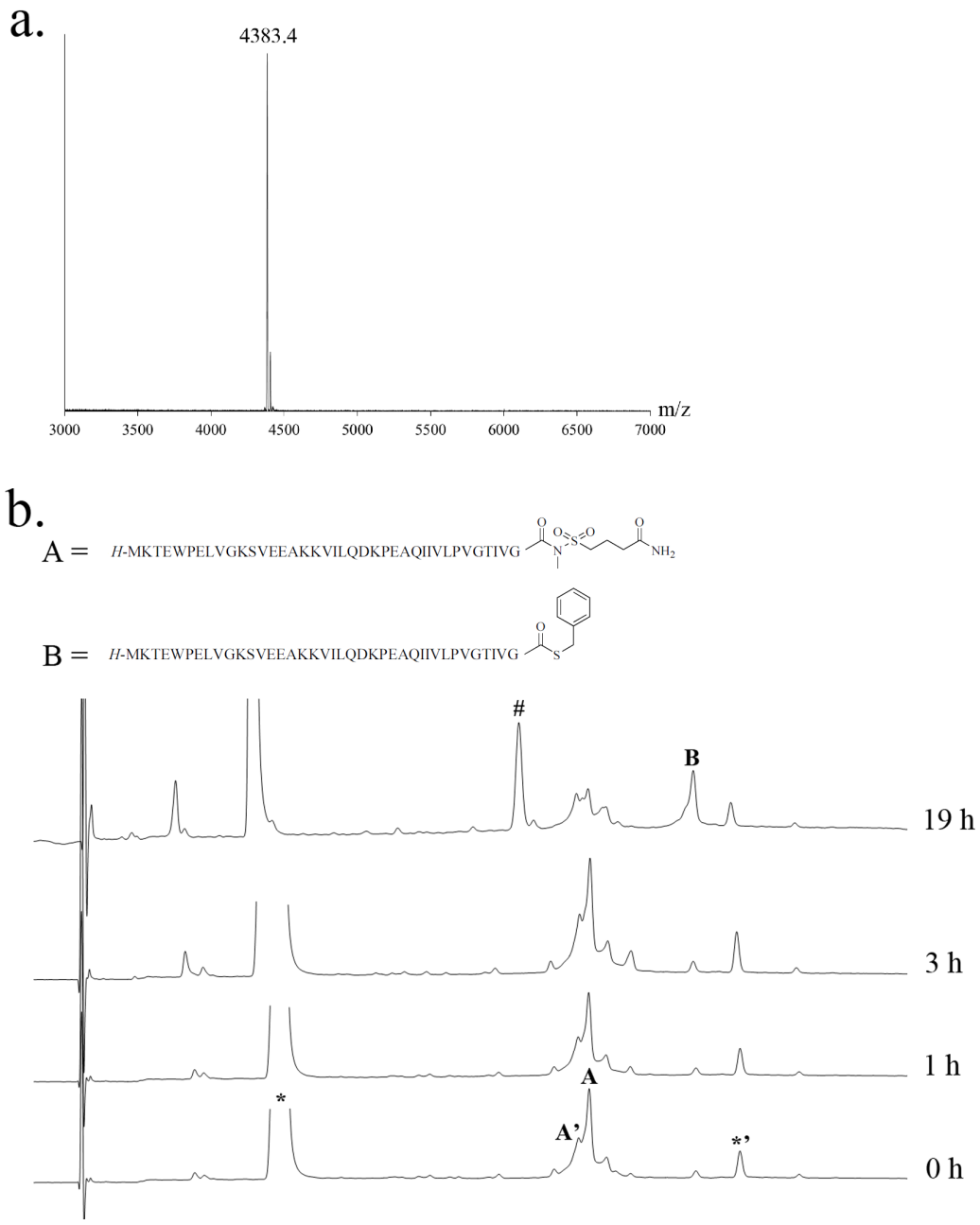


Figure 5.2: CI2(1-39)T39G N-(Me)sulfonamide thioester exchange with MPAA and benzyl mercaptan

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 4383.4 Da, calculated product mass = 4380.22 Da.

b. Analytical-HPLC time course of exchange. Exchange conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCL (50 mM), MPAA (60 mM), benzyl mercaptan (10 % v/v from 4 h), pH 7.0. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(1-39)T39G N-(Me)sulfonamide, A' = CI2(1-39)T39G impurity, B = CI2(1-39)T39G benzyl mercaptan thioester, * = MPAA (peak truncated), *' = MPAA impurity, # = benzyl mercaptan.

5.2.4 Thioester exchange with MPAA and benzyl mercaptan at 40 °C

The conditions with both thiols were replicated but the temperature increased to 40 °C to increase the rate, and to have both thiols present from t = 0.

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.44 mg) was resuspended in non-denaturing buffer with TCEP-HCL (50 mM), benzyl mercaptan (10 % v/v) and MPAA (60 mM), pH 7.0. The reaction was performed at 40 °C with 350 rpm shaking and monitored by analytical-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. The exchange is now visibly quicker with a substantial amount of product by 7 h, collected and characterized by MALDI-TOF analysis: observed mass = 4381.9 Da in positive reflector mode, calculated mass = 4380.22 Da (Figure 5.3).

This was then repeated on a larger scale to collect CI2(1-39)T39G benzyl mercaptan thioester to perform conformationally-assisted ligation with a pre-exchanged thioester peptide. CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 34 mg) dissolved in non-denaturing buffer with TCEP-HCL (50 mM), MPAA (60 mM) and benzyl mercaptan (10 % v/final v), pH 7.0. Monitored at t = 0 and again after 8 h by analytical-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. At this time CI2(1-39)T39G benzyl mercaptan thioester purified by semi preparative-HPLC: RP-C18 column with a gradient of 30 – 70 % B in 30 min and a flow rate of 5 ml min^{-1}. Purified peptide freeze-dried (8 mg, 23.5 % yield).

5.2.5 Ligation with MPAA and benzyl mercaptan

Returning to conformationally-assisted ligation, the successful thioester exchange conditions were used. CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.46 mg)

was mixed with CI2(40-64) (1 mM, 0.42 mg) in non-denaturing buffer with TCEP-HCl (50 mM), MPAA (60 mM) and benzyl mercaptan (10 %), pH 6.3. The reaction was monitored by analytical-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. Over time the CI2(1-39)T39G benzyl mercaptan thioester forms, but no ligation product (data not shown). The intermediate was collected and characterized by MALDI-TOF analysis: observed mass = 4383.2 Da in positive reflector mode, calculated mass = 4380.22 Da.

5.2.6 Ligation with pre-formed thioester peptide and thiophenol

With CI2(1-39)T39G as a thioester peptide, the published conditions of thiophenol (2 % v/v) (Beligere and Dawson, 1999) were used to attempt conformationally-assisted ligation with a pre-exchanged thioester peptide.

CI2(1-39)T39G benzyl mercaptan thioester (1 mM, 0.22 mg) was mixed with CI2(40-64) (1 mM, 0.15 mg) in non-denaturing buffer with thiophenol (2 % v/v), pH 7.5. The reaction was monitored by analytical-HPLC: RP-C18 column with a gradient of 30 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. The thiophenol was added to the stock buffer and so 1 µl extra was added directly to the ligation reaction at 3 h in case not enough had been transferred. Over time observed the formation of CI2(1-39)T39G thiophenol thioester, but no ligation product (Figure 5.4). The thiophenol thioester was collected and confirmed by MALDI-TOF analysis: observed mass = 4369.2 Da in positive reflector mode, calculated mass = 4366.19 Da.

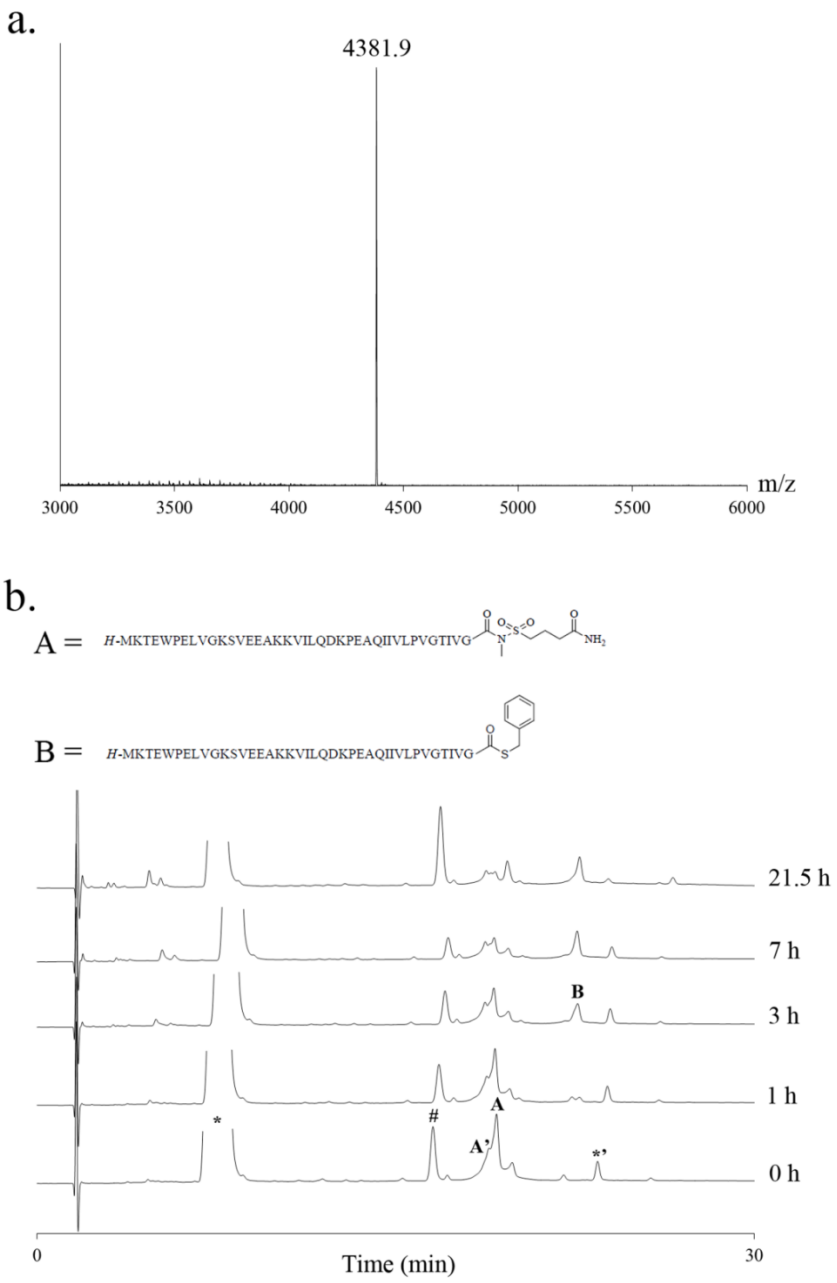


Figure 5.3: CI2(1-39)T39G N-(Me)sulfonamide thioester exchange with MPAA and benzyl mercaptan at 40 °C

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 4381.9 Da, calculated product mass = 4380.22 Da.

b. Analytical-HPLC time course of exchange. Exchange conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (60 mM), benzyl mercaptan (10 % v/v), pH 7.0, 40 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(1-39)T39G N-(Me)sulfonamide, A' = CI2(1-39)T39G impurity, B = CI2(1-39)T39G benzyl mercaptan thioester, * = MPAA (peak truncated), *' = MPAA impurity, # = benzyl mercaptan.

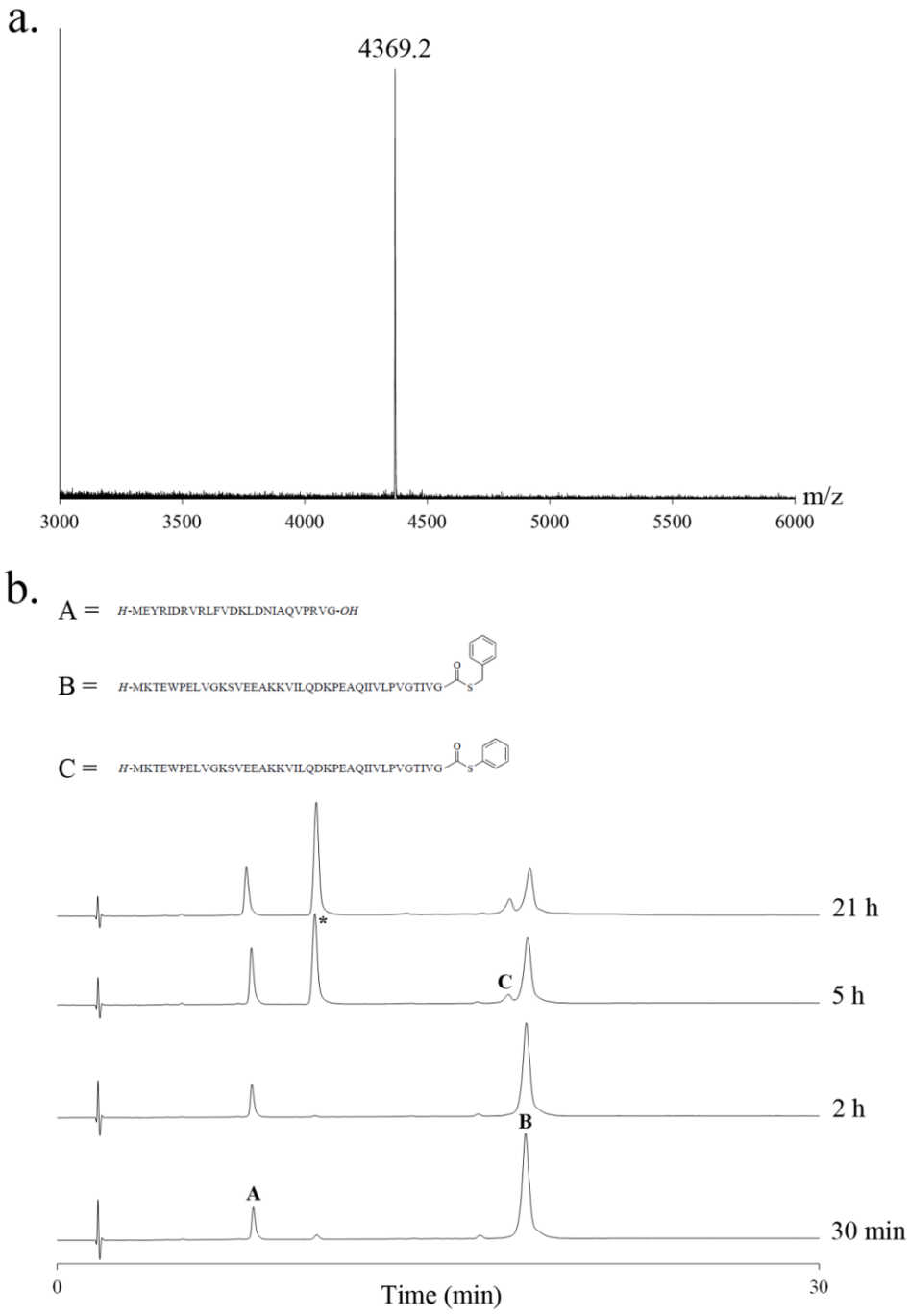


Figure 5.4: Conformationally-assisted ligation of CI2(1-39)T39G benzyl mercaptan thioester with CI2(40-64)

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 4369.2 Da, calculated product mass = 4366.19 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptide (1 mM), sodium phosphate (100 mM), thiophenol (2 % v/v), pH 7.5. HPLC conditions: RP-C18 column, 30 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(40-64), B = CI2(1-39)T39G benzyl mercaptan thioester, C = CI2(1-39)T39G thiophenol thioester.

5.3 Biophysical characterization

Attempts to ligate with a Gly residue at the interaction site yielded no success, so the interaction was further investigated. After the success of BLI with the RNase S peptides this was the initial route of examination, followed by fluorescence spectroscopy. CI2(1-39)T39G was biotinylated at its *N*-terminus for immobilization to streptavidin-coated sensors, to monitor the interaction at its *C*-terminus with CI2(40-64). The concentrations of the peptides were established by UV/VIS spectrophotometry.

5.3.1 Preparation of peptides

Biotinylated CI2(1-39)T39G

Biotinylated CI2(1-39)T39G was synthesized on pre-loaded Fmoc-Gly-Wang resin (0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Upon completion of synthesis biotin was coupled at the *N*-terminus, confirmed with a negative Kaiser test. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 20 – 70 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (14 mg, 2.76 % yield).

5.3.2 BLI with CI2(1-39)T39G

Biotin-CI2(1-39)T39G (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *N*-terminus and is introduced to varying concentrations of CI2(40-64) (20 – 100 nM) in a buffer of HEPES (10 mM), Tween (0.005 %) and EDTA (2 mM), pH 7.0. Unable to determine a value for the K_d after no interaction was observed.

5.3.3 Fluorescence spectroscopy

After BLI proved inappropriate for surveying affinity, fluorescence spectroscopy was used. If the peptides are interacting, a decrease in fluorescence should be observed as Trp5 and Tyr42 become obscured when the two peptides interact and fold.

Fluorescence data were obtained by titrating varying concentrations of CI2(40-64) (1.88, 2.81 and 4.65 µM) into CI2(1-39)T39G (2.3 µM), and after 25 s the fluorescence at 345 nm monitored over 800 s. All peptides in a PBS buffer at pH 7.4. Decreased fluorescence and a concentration-dependent rate establishes a high affinity interaction is taking place (Figure 5.5).

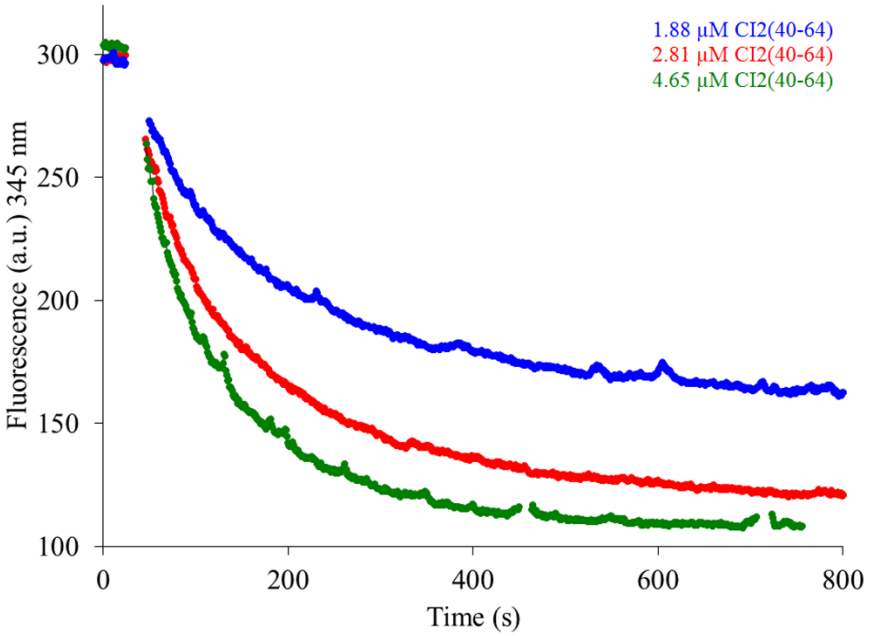


Figure 5.5: Fluorescence spectroscopy of CI2 peptide interaction

CI2(40-64) (1.88, 2.81 and 4.65 µM) is titrated into CI2(1-39)T39G (2.3 µM), and after 25 s the fluorescence at 345 nm monitored over 800 s, in PBS buffer at pH 7.4.

5.4 Conformationally-assisted ligation of CI2(1-39)T39G N-(Me) sulfonamide with CI2(40-64)M40C

Although the published work detailed in section 5.1 by Beligere and Dawson demonstrated conformationally-assisted chemical ligation with the native *N*-terminal Met of CI2(40-64), the rate was significantly improved with its substitution for Cys, supporting the idea that peptide conformation can be exploited to encourage ligation at difficult sites. CI2(40-64)M40C was synthesized to encourage ligation with CI2(1-39)T39G N-(Me)sulfonamide.

5.4.1 Preparation of peptides

CI2(40-64)M40C

CI2(40-64)M40C was synthesized on pre-loaded Fmoc-Gly-Wang resin (0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 30 % in 5 min, 30 – 50 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (21.4 mg, 5.85 % yield).

5.4.2 Ligation with thiophenol

This first ligation attempt used the thiol in the paper discussed in 5.1, thiophenol, and was carried out under folding conditions.

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.66 mg) added to CI2(40-64)M40C (1 mM, 0.47 mg) in degassed non-denaturing buffer with TCEP-HCl (50 mM) and thiophenol (20 % v/v), pH 6.3. The reaction was monitored by analytical-

HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. 20 % thiophenol was accidentally used rather than 2 %, but under these conditions the ligation was complete in 2 h (Figure 5.6). The ligation product was collected for MALDI-TOF confirmation: observed mass = 7238.0 Da in positive reflector mode, calculated mass = 7231.51 Da.

5.4.3 Ligation with thiophenol, denaturing conditions

To demonstrate that the ligation in section 5.4.2 is conformationally-assisted, a control reaction using denaturing conditions was carried out.

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.29 mg) added to CI2(40-64)M40C (1 mM, 0.20 mg) in degassed denaturing buffer with TCEP-HCl (50 mM) and thiophenol (2 % v/v), pH 6.3. The reaction monitored by analytical-HPLC: RP-C18 column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. Though 2 % thiophenol was used in this instance, no ligation is observed by 24 h (Figure 5.7).

5.4.4 Ligation with MPAA

Its smell and solubility do not make thiophenol the ideal choice for a ligation thiol. As such the final ligation used MPAA.

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.29 mg) added to CI2(40-64)M40C (1 mM, 0.20 mg) in degassed non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 6.5. The reaction monitored by analytical-HPLC: RP-C4

column with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. A C4 column was chosen because of the poor separation of CI2(40-64)M40C and thiophenol in the previous section. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 5.8). Conformationally-assisted ligation proceeds to completion in 24 h, formed even in the t = 0 time point's HPLC trace. The product was collected and characterized by MALDI-TOF analysis: observed mass = 7234.0 Da in positive linear mode, calculated mass = 7231.51 Da.

Conformationally-assisted ligation with MPAA as the chosen thiol was repeated to deduce a yield for the reaction, taking time points at just t = 0 and the end of the reaction rather than monitoring throughout.

CI2(1-39)T39G N-(Me)sulfonamide (1 mM, 0.30 mg) was mixed with CI2(40-64)M40C (1 mM, 0.22 mg) in degassed non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 6.5. Time points were taken at t = 0 and 26 h and checked by analytical-HPLC: RP-C4 column with a gradient of 20 – 70 % B in 30 min and a flow rate of 1 ml min^{-1}. Confirmed to be complete, the remaining volume was injected over four runs to collect the ligation product. Purified peptide freeze-dried (0.5 mg, 96.2 % yield).

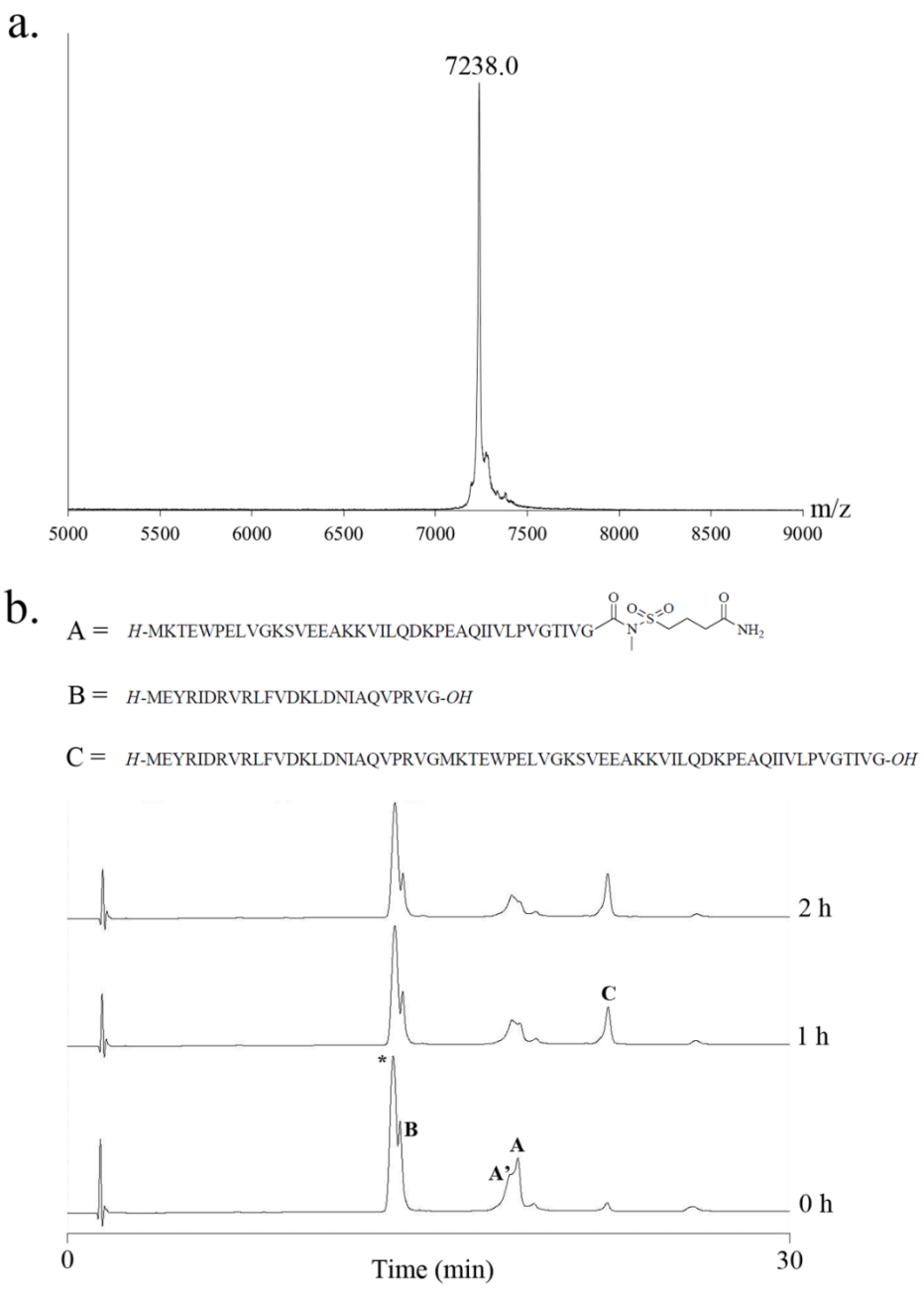


Figure 5.6: Conformationally-assisted ligation of CI2(1-39)T39G N-(Me)sulfonamide with CI2(40-64)M40C, with thiophenol

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 7238.0 Da, calculated product mass = 7231.51 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), thiophenol (20 % v/v), pH 6.3. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(1-39)T39G N-(Me)sulfonamide, A' = CI2(1-39)T39G impurity, B = CI2(40-64)M40C, C = ligation product * = thiophenol.

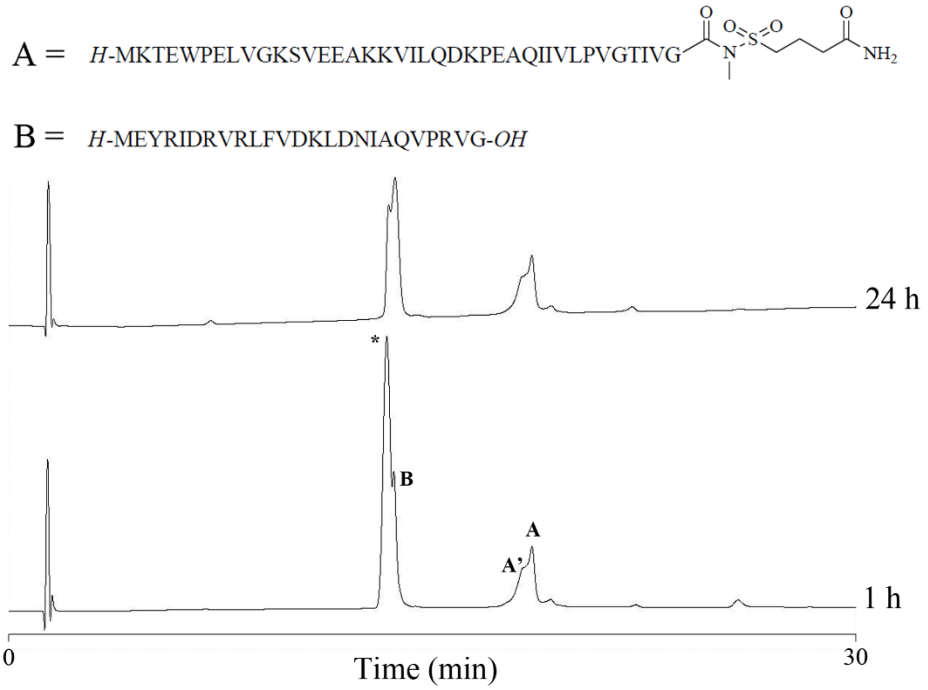


Figure 5.7: Conformationally-assisted ligation of CI2(1-39)T39G N-(Me)sulfonamide with CI2(40-64)M40C, denaturing conditions

Analytical-HPLC time course of ligation. Ligation conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), thiophenol (2 % v/v), pH 6.3. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(1-39)T39G N-(Me)sulfonamide, A' = CI2(1-39)T39G impurity, B = CI2(40-64)M40C.

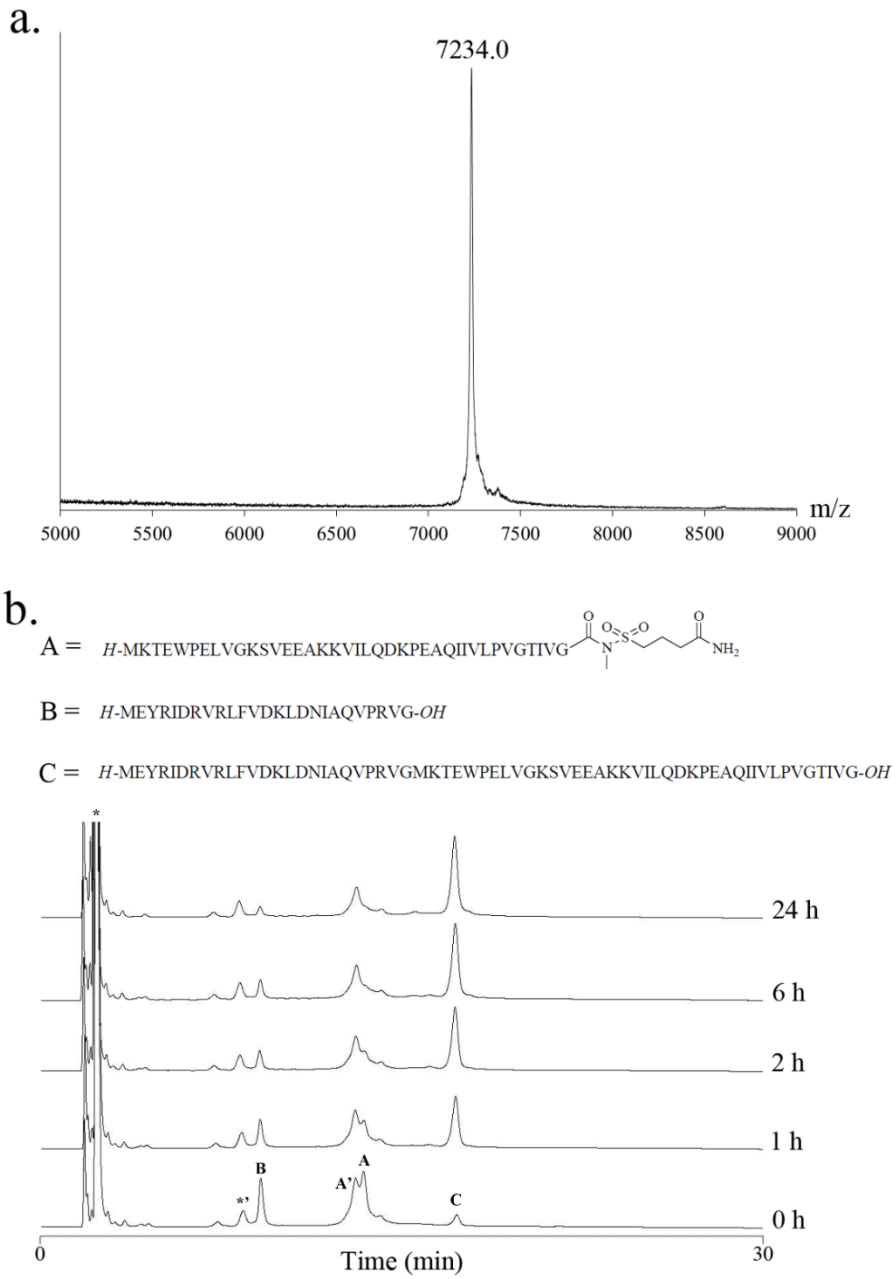


Figure 5.8: Conformationally-assisted ligation of CI2(1-39)T39G N-(Me)sulfonamide with CI2(40-64)M40C, with MPAA

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 7234.0 Da, calculated product mass = 7231.51 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (60 mM), pH 6.5. HPLC conditions: RP-C4 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = CI2(1-39)T39G N-(Me)sulfonamide, A' = CI2(1-39)T39G impurity, B = CI2(40-64)M40C, C = ligation product * = MPAA (peak truncated).

5.5 Attempting conformationally-assisted ligation with CI2

5.5.1 Conformationally-assisted ligation

The idea of using CI2 to demonstrate conformationally-assisted ligation using a peptide N-(Me)sulfonamide was based upon the published work discussed at the beginning of this chapter, where CI2 peptide fragments were shown to ligate under folding conditions in both the presence and absence of a Cys at the ligation site (Beligere and Dawson, 1999). In this paper the native Thr at the *C*-terminus of CI2(1-39) is replaced with Asp, so we hypothesized that replacing it with Gly would have no detrimental effect. A Gly at the ligation site should increase the opportunity for conformationally-assisted ligation by minimising steric hindrance at the reaction site.

This first set of ligations used CI2(1-39)T39G N-(Me)sulfonamide and CI2(40-64), with its native *N*-terminal Met. The initial attempt used folding conditions and 60 mM MPAA as the thiol additive, at pH 6.3 as with the work of Beligere and Dawson. The thioester exchange intermediate, CI2(1-39)T39G MPAA thioester, was observed but no ligation product (Figure 5.1). It was decided to raise the pH in line with the previous peptide N-(Me)sulfonamide ligations and so a replica experiment was conducted at pH 7.5. This too generated CI2(1-39)T39G MPAA thioester but no ligation product.

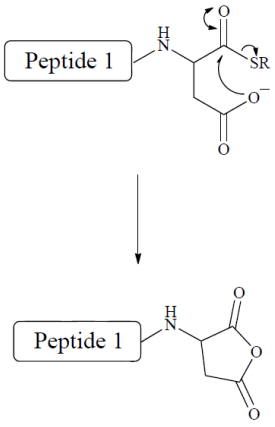
As the thioester exchange step readily occurs under these conditions it was decided to define the ideal conditions for this to take place, in order to collect enough to ligate with the pre-exchanged thioester peptide. Firstly the above ligation conditions were repeated but without the presence of CI2(40-64), and this exchange with MPAA was relatively slow. In order to leave the reaction overnight benzyl mercaptan was added with the thought that CI2(1-39)T39G benzyl mercaptan thioester might be more stable to this lengthy exchange, benzyl mercaptan being a strong nucleophile and thus

providing a less reactive thioester peptide. The time point of the following morning showed near complete conversion to the benzyl mercaptan thioester, and no remaining MPAA thioester (Figure 5.2).

Having observed this conversion the next reaction left out the MPAA and used benzyl mercaptan (10 % v/v) alone. Though the benzyl mercaptan thioester forms it does so slower than when MPAA is present, and so the MPAA is acting as a catalyst to this reaction. This combination of MPAA and benzyl mercaptan was re-employed but with increased temperature, the reaction now at 40 °C (Figure 5.3). There is substantial exchange product present by the 7 h time point, and so the reaction was repeated on a larger scale for collection of CI2(1-39)T39G benzyl mercaptan thioester to perform ligation with a pre-exchanged thioester peptide.

To conduct a thorough investigation of conformationally-assisted ligation between CI2(1-39)T39G N-(Me)sulfonamide and CI2(40-64), the successful exchange conditions combining MPAA and benzyl mercaptan were used. No ligation product is observed, despite the benzyl mercaptan thioester peptide readily forming under these conditions. Finally the pre-exchanged CI2(1-39)T39G benzyl mercaptan thioester was added to CI2(40-64) under conditions similar to those published, in non-denaturing buffer and with 2 % thiophenol (Figure 5.4). Due to the nature of thiophenol's solubility an additional 1 µl was added directly to the reaction mix after 3 h, in case the lack of ligation was a result of not enough thiophenol transferred from the stock buffer. The CI2(1-39)T39G benzyl mercaptan thioester converts to the thiophenol thioester, which is more weakly activated and therefore less likely to ligate. Future work could involve using a pre-exchanged thioester peptide without any thiol additives, though no ligation was observed prior to the addition of extra thiophenol in this most recent attempt.

We wondered if the use of CI2(1-39)T39D was a true representation of conformationally-assisted ligation, or if an anhydride intermediate played a role. Since the publication of the ligation by Beligere and Dawson, Botti and co-workers identified the formation of a succinic anhydride intermediate of Asp-Cys ligation formed during attack by the Asp side chain (Villain et al., 2003) (Scheme 5.1). This reaction is later discussed by Kent and colleagues who confirmed the unsuitability of an Asp-Cys site for ligation (Dang et al., 2013). The succinic anhydride is more highly activated than the original thioester peptide, and may have been responsible for the increased observed rate of the published conformationally-assisted ligation. Future work could include the synthesis of CI2(1-39)T39D to try to replicate the Beligere and Dawson work, but even if we did the lack of selectivity provided by the anhydride would not be ideal for a protein labelling method, as we could not guarantee that coupling to other *N*-termini besides that targeted would not occur.



Scheme 5.1: Succinic anhydride formation at a *C*-terminal Asp thioester peptide

Attack of the *C*-terminal Asp side chain in the thioester peptide can form the succinic anhydride intermediate (Villain et al., 2003, Dang et al., 2013) that is more highly activated than the original thioester peptide, resulting in increased ligation rates but reduced selectivity.

5.5.2 Biophysical characterization

After no conformationally-assisted ligation was observed between CI2(1-39)T39G N-(Me)sulfonamide and CI2(40-64) the interaction between these peptides was investigated further with biophysical techniques. The native Thr at the *C*-terminus of CI2(1-39) has been replaced by Gly, and though this was meant to minimise steric hindrance and promote ligation perhaps it has detrimentally affected the affinity between the peptides.

When the RNase A peptides were studied in this way, BLI was used to establish K_d values. CI2(1-39)T39G was synthesized with *N*-terminal biotin for immobilization to streptavidin-coated sensors, maintaining a free *C*-terminus for interaction with CI2(40-64). Surprisingly no interaction was observed using this technique. One possibility could be if CI2(1-39)T39G is too close to the sensor for CI2(40-64) to access its *C*-terminus, and to test this a linker could be introduced at its *N*-terminus.

Concerned that these peptides might not be interacting, fluorescence spectroscopy was used to probe further. Both peptides have aromatic residues capable of emitting fluorescence when excited, CI2(1-39)T39G with Trp5 and CI2(40-64) with Tyr42. If the peptides are interacting the emitted fluorescence should decrease, where these residues become buried upon peptide folding. CI2(40-64) was titrated into CI2(1-39)T39G at varying concentrations, and after 25 s their interaction monitored at 345 nm for 800 s (Figure 5.5). A concentration-dependent interaction is observed, and the decrease in fluorescence over this time tells us that the peptides are definitely interacting and folding. Though this is promising in terms of their affinity for one another and theoretical capability for conformationally-assisted ligation, it does not help to explain why this has not yet taken place. As with RNase A no conditions besides that

of the ligation are provided in the literature for conformationally-assisted ligation with CI2, and so there were no obvious ideas for improvement.

5.5.3 Conformationally-assisted ligation with Cys

Having established that a non-covalent interaction is taking place at this reaction site CI2(40-64)M40C was synthesized for renewed conformationally-assisted ligation attempts. Though ligation at a Gly-Met site was not demonstrated the fluorescence spectroscopy results show that interaction at this site is happening, and so introducing a Cys to this site should allow for ligation. Conformationally-assisted ligation with an Asp-Cys at the ligation site takes a quarter of the time as with the native Met (Beligere and Dawson, 1999), and so our Gly-Cys ligation should stand a better chance of working than the Gly-Met site investigated in section 5.2.

The first attempt between CI2(1-39)T39G N-(Me)sulfonamide and CI2(40-64)M40C took place under the published conditions, phosphate buffer at pH 6.3 with thiophenol. 20 % thiophenol was used in place of the 2 % in literature, and under these conditions ligation was complete in 2 h (Figure 5.6). The addition of Cys has resulted in a fast ligation, and the thioester intermediate is not observed.

To show that the ligation is conformationally-assisted it was next conducted under denaturing conditions. Some ligation may be expected, as we are investigating a Gly-Cys ligation site, but without peptide conformation any ligation should be significantly slower. No ligation is observed and so it can be concluded that the previous ligation is in fact conformationally-assisted, with the structure of the peptides crucial to its success. Though 2 % thiophenol was used instead of 20 % in this instance there is still no evidence of ligation after 24 h (Figure 5.7).

The offensive odour of thiophenol makes it not the favourite choice for ligation reactions. As discussed in Chapter 4 MPAA is commonly used as a thiol additive, and so the current ligation was repeated with MPAA. Ligation was complete in 24 h and the thioester intermediate is not observed, with ligation product visible from as early as the t = 0 trace (Figure 5.8).

5.6 Summary

Following the ligation problems with RNase A, CI2 was chosen as the next split protein system. CI2(1-39)T39G N-(Me)sulfonamide was synthesized for conformationally-assisted ligation studies with CI2(40-64). No ligation is detected, even with the pre-formed thioester, but fluorescence spectroscopy confirmed that the peptides were interacting and folding. CI2(40-64)M40C was synthesized and successfully ligated to CI2(1-39)T39G N-(Me)sulfonamide under folding conditions. The search continues for a split protein capable of conformationally-assisted ligation at a Cys-free site.

6. Conformationally-assisted reactions with calbindin D_{9k}

6.1 Introduction

The work described so far on the limited success with both RNase A and CI2 has led to the search for a third split protein. Following the work with RNase A, CI2 was chosen for, among other reasons, its greater affinity. No conformationally-assisted ligation was observed in this split protein unless a Cys residue was introduced to the reaction site, and so we wanted to find a split protein with fragments possessing a higher affinity again. A split protein that stood out for us, noted from the Linse review upon which Table 1.2 is based (Carey et al., 2007), was a calbindin D_{9k} variant. A very high affinity is reported between its two peptides, an interaction with a K_d of 3 pM (Berggård et al., 2001), in addition to the relatively small size of both peptides. The work published in this paper investigated the contributing factors of non-covalent interactions on protein structure, notably the role of residue substitutions in the hydrophobic core. It was demonstrated in reconstitution studies that the hydrophobic core of a protein plays a key role in the generation of fragment complexes with native structures and activities (Richards, 1958), and so SPR was employed to study the effect of introducing mutations in this area. Though dissociation curves were not observed as concentration-dependent, association curves were: the equilibrium association constant K_A for wt-EF1 was determined as 3.5×10^{11} M^{-1}. We thought that two peptides with such an affinity could surely undergo conformationally-assisted ligation, and so calbindin D_{9k} was chosen as our final split protein to study.

6.2 Biophysical characterization

Fluorescence spectroscopy was also used by Linse and co-workers to confirm folding and complex formation, using an EF2 mutant with a F66W substitution. It was decided to use BLI to study the interaction between our synthetic EF1 and EF2 peptides, following successful K_d determination for RNase A in Chapter 4. The F66W variant was used rather than wt-EF2 to allow for concentration determination by UV/VIS spectrophotometry, and in each experiment the EF2 peptide was immobilized to a streptavidin-coated sensor *via* its biotinylated *C*-terminus. Three interaction sites were investigated by substituting the amino acids at the appropriate *N*- of *C*-termini. The first interaction was between wt-EF1 and EF2 F66W, with a Gly-Met interaction. The second used a Ser-Ser interaction, and the final investigated a Gly-Gly interaction.

Following this CD spectroscopy was used to probe the extent of conformation held by the peptides possessing a Gly-Gly interaction site, and the effect of calcium on this conformation.

6.2.1 Preparation of peptides

EF1

EF1 synthesized on pre-loaded Fmoc-Gly-Wang resin (0.76 mmol g^{-1} loading, 0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in the denaturing buffer with TCEP-HCl (50 mM) and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 30 % B in 5 min, 30 – 70 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (68.3 mg, 12.1 % yield).

EF1+S

EF1+S synthesized on pre-loaded Fmoc-Ser(tBu)-Wang resin (0.05 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in the denaturing buffer with TCEP-HCl (50 mM) and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 30 % B in 5 min, 30 – 70 % B in 30 min and a flow rate of 15ml min^{-1}. Purified peptide freeze-dried (64.8 mg, 22.6 % yield).

EF2 F66W-PEG biotin

EF2 F66W-PEG biotin was synthesized on Fmoc-PEG biotin NovaTag resin (0.36 mmol g^{-1} loading, 0.05 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in the denaturing buffer with TCEP-HCl (50 mM) and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 35 % B in 5 min, 35 – 70 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (6 mg, 2.54 % yield).

EF2 ΔM43/F66W-PEG biotin

EF2 ΔM43/F66W-PEG biotin was synthesized on Fmoc-PEG biotin NovaTag resin (0.36 mmol g^{-1} loading, 0.075 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in the denaturing buffer with TCEP-HCl (50 mM) and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 35 % B in 5 min, 35 – 70 % in 30 min and a flow rate of 15ml min^{-1}. Purified peptide freeze-dried (2.1 mg, 0.6 % yield).

EF2 ΔM43/S44G/F66W-PEG biotin

EF2 ΔM43/S44G/F66W-PEG biotin was synthesized on Fmoc-PEG biotin NovaTag resin (0.36 mmol g^{-1} loading, 0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in the denaturing buffer with TCEP-HCl (50 mM) and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 35 % B in 5 min, 35 – 70 % in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (32.8 mg, 7.18 % yield).

EF2 ΔM43/S44GF66W-NH_2

EF2 ΔM43/S44G/F66W-NH_2 was synthesized on PL rink resin (0.39 mmol g^{-1} loading, 0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in the denaturing buffer and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 35 % B in 5 min, 35 – 70 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (yield not determined).

6.2.2 BLI of Gly-Met interaction

EF2 F66W-PEG biotin (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *C*-terminus and is introduced to varying concentrations of EF1 (0.1 – 10 μM), both peptides in a buffer of HEPES (10 mM), Tween (0.005 %) and $CaCl_2$ (2 mM) at pH 7.0. A K_d value of 47.7 nM ± 5.8 nM was observed for [EF1] 0.1 – 4 μM (Figure 6.1).

6.2.3 BLI of Ser-Ser interaction

EF2 ΔM43/F66W-PEG biotin (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *C*-terminus and is introduced to varying concentrations of EF1+S (0.1 – 10 μM), both peptides in a buffer of HEPES (10 mM), Tween (0.005 %) and $CaCl_2$ (2 mM) at pH 7.0. A K_d value of 59.8 nM ± 12.8 nM was observed for [EF1+S] 0.1 – 4 μM (Figure 6.2).

6.2.4 BLI of Gly-Gly interaction

EF2 ΔM43/S44G/F66W-PEG biotin (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *C*-terminus and is introduced to varying concentrations of EF1 (0.1 – 10 μM), both peptides in a buffer of HEPES (10 mM), Tween (0.005 %) and $CaCl_2$ (2 mM) at pH 7.0. A K_d value of 69.0 nM ± 9.8 nM was observed for [EF1] 0.1 – 10 μM (Figure 6.3).

6.2.5 BLI of Gly-Gly interaction in the presence of EDTA

The above BLI was repeated but in a buffer containing EDTA in place of calcium, to begin to establish the effect of calcium on the peptides' interaction.

EF2 ΔM43/S44G/F66W-PEG biotin (2 μg ml^{-1}) binds streptavidin-coated sensors *via* its *C*-terminus and is introduced to varying concentrations of EF1 (0.1 – 10 μM), both peptides in a buffer of HEPES (10 mM), Tween (0.005 %) and EDTA (2 mM) at pH 7.0. The raw data indicated non-specific binding to the sensor (Figure 6.4).

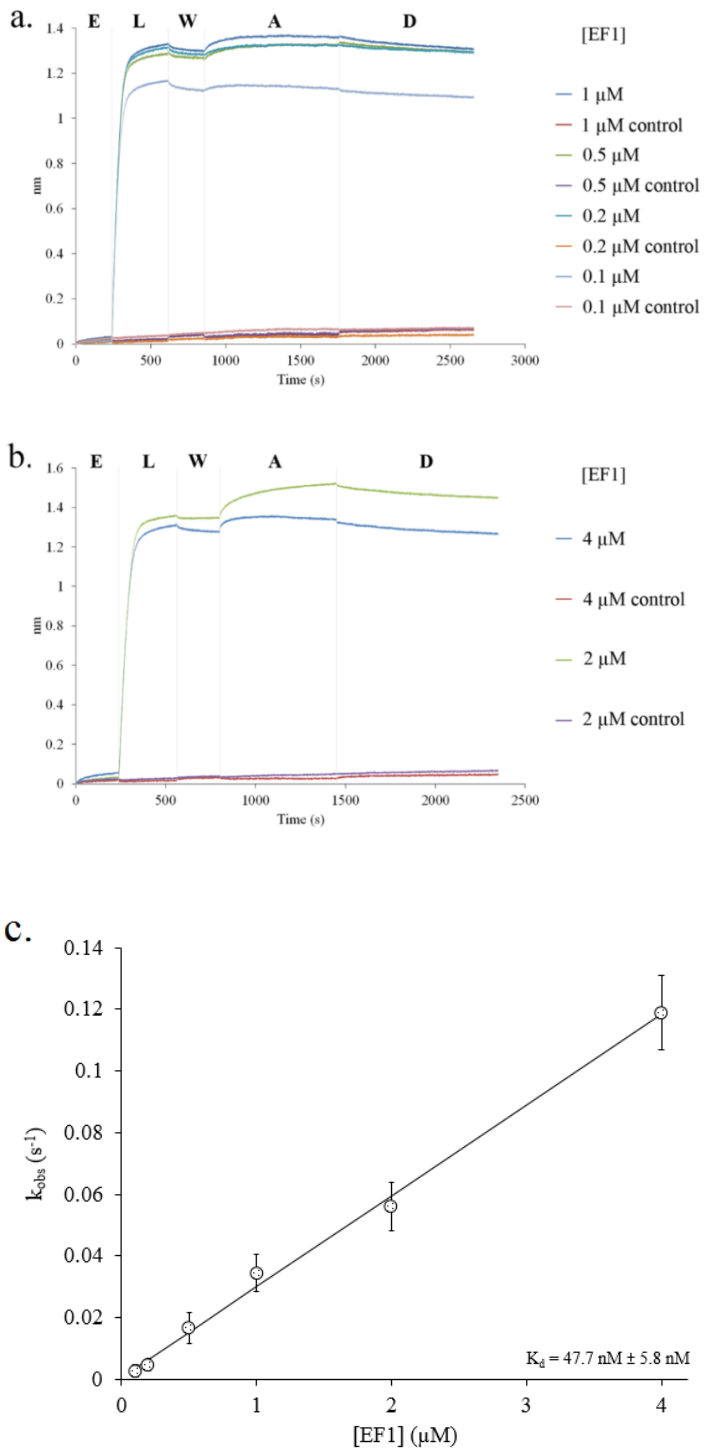


Figure 6.1: BLI analysis of Gly-Met interaction

a. Raw data for [EF1] 0.1 – 1 µM interacting with immobilized EF2, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation.

b. Raw data for [EF1] 2 – 4 µM interacting with immobilized EF2, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation.

c. k_{obs} plotted against [EF1], the gradient giving a K_d of 47.7 nM ± 5.8 nM.

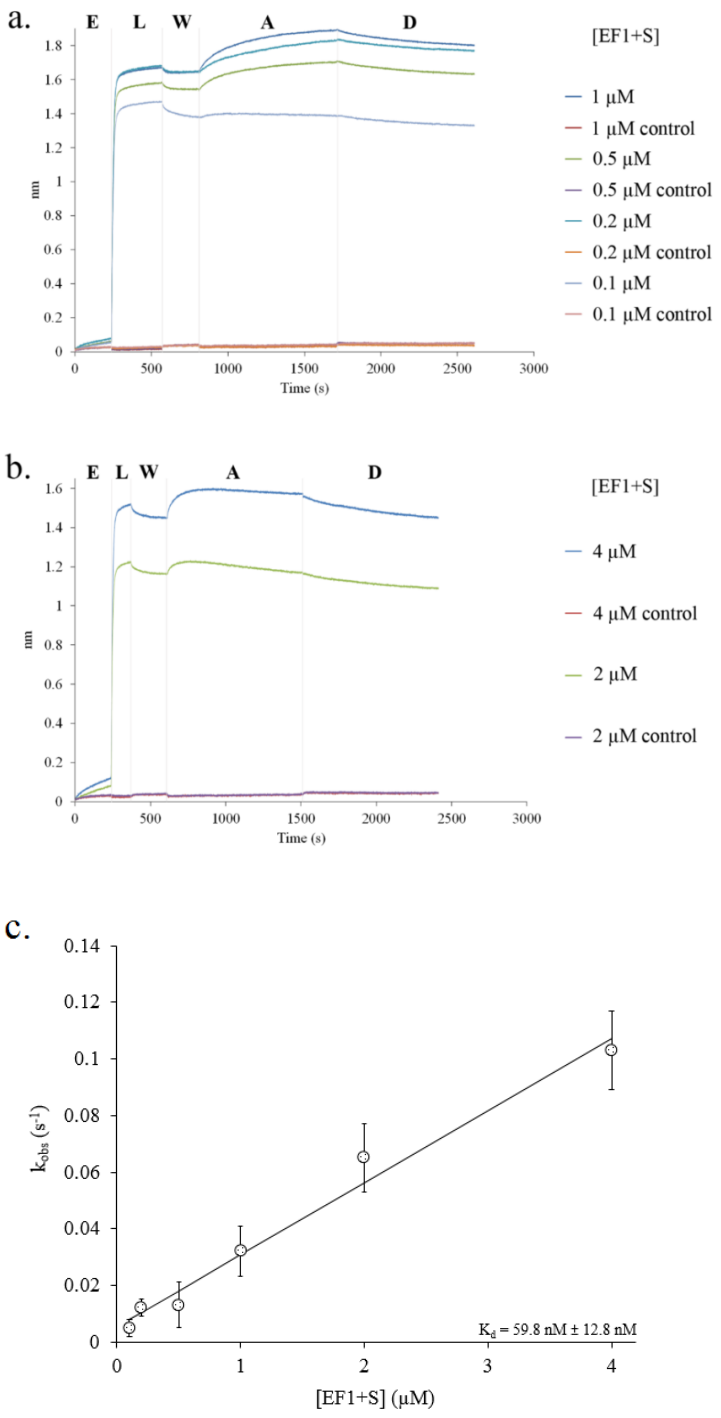


Figure 6.2: BLI analysis of Ser-Ser interaction

a. Raw data for [EF1+S] 0.1 – 1 µM interacting with immobilized EF2, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation.

b. Raw data for [EF1+S] 2 – 4 µM interacting with immobilized EF2, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation.

c. k_{obs} plotted against [EF1+S], the gradient giving a K_d of 59.8 nM ± 12.8 nM.

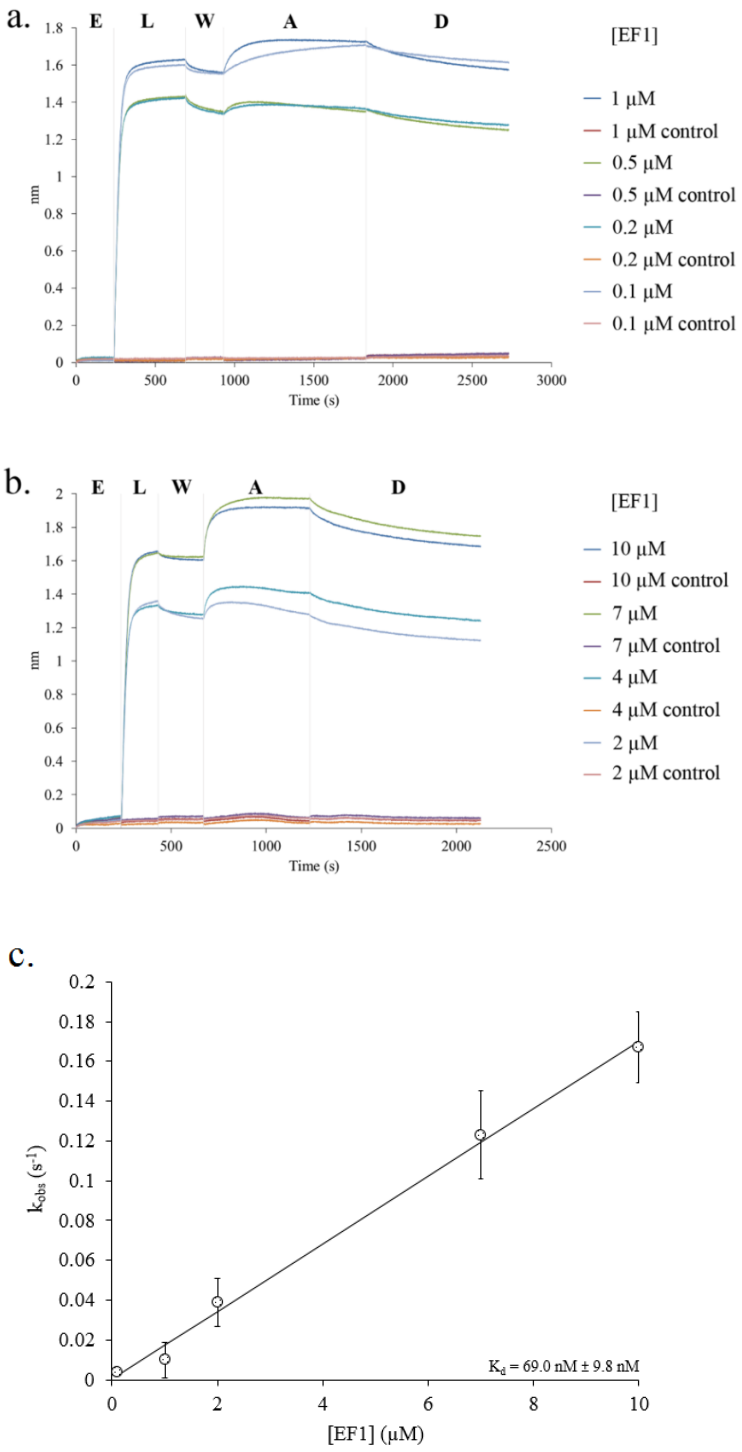


Figure 6.3: BLI analysis of Gly-Gly interaction

a. Raw data for [EF1] 0.1 – 1 µM interacting with immobilized EF2, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation.

b. Raw data for [EF1] 2 – 10 µM interacting with immobilized EF2, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation.

c. k_{obs} plotted against [EF1], the gradient giving a K_d of 69.0 nM ± 9.8 nM.

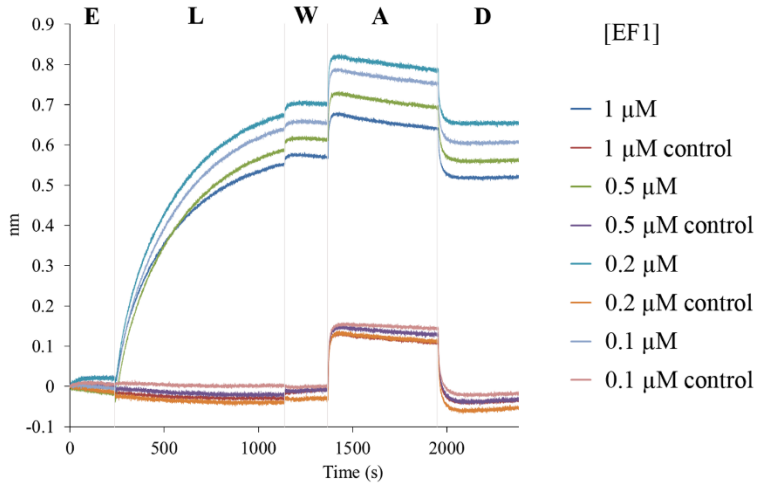


Figure 6.4: BLI analysis of Gly-Gly interaction with EDTA

Raw data for [EF1] 0.1 – 1 µM interacting with immobilized EF2 in the presence of a chelator, along with controls with EF2 blank replacements. E = equilibration, L = loading, W = wash, A = association, D = dissociation. Non-specific binding is observed, illustrated by the response of the control experiments.

6.2.6 CD spectroscopy of Gly-Gly interaction

The next line of enquiry was to determine the level of folding under 2 mM calcium, and compare it to in the presence of a calcium chelator. CD experiments were conducted with both EF1 and EF2 ΔM43/S44G/F66W-NH_2 in either the CD EDTA or calcium buffer, first individually and then when mixed at low micromolar concentrations. In the EDTA buffer EF1 (30.4 µM) and EF2 S44G/F66W-NH_2 (25.2 µM) were first analysed individually and then after mixing, in a 2 mm cuvette. In the calcium buffer EF1 (47.4 µM) and EF2 S44G/F66W-NH_2 (41.5 µM) were again analysed separately before mixing, in a 1 mm cuvette (Figure 6.5). Wavelengths from 200 – 250 nm were measured in both experiments.

In each buffer both EF1 and EF2 ΔM43/S44G/F66W-NH_2 showed α–helical content (Table 6.1), and in the calcium buffer the interacting peptides had greater α-helical content than that predicted. EF1 was found to contain 17.7 % α-helix in the EDTA-containing buffer in comparison with 43.8 % in the calcium-containing buffer, and EF2 was shown to consist of 19.8 % α-helix in the absence compared to 37.2 % with 2 mM calcium.

	%			
	α-helix	**β-sheet**	**Turn**	**Random**
EF1 with EDTA	17.7	16.8	14.6	51.0
EF2 with EDTA	19.8	14.6	14.7	50.9
EF1 with calcium	43.8	11.8	16.0	28.4
EF2 with calcium	37.2	16.6	18.2	28.0

Table 6.1: The structural content of calbindin D_{9k} peptides with 2 mM $CaCl_2$ or 2 mM EDTA

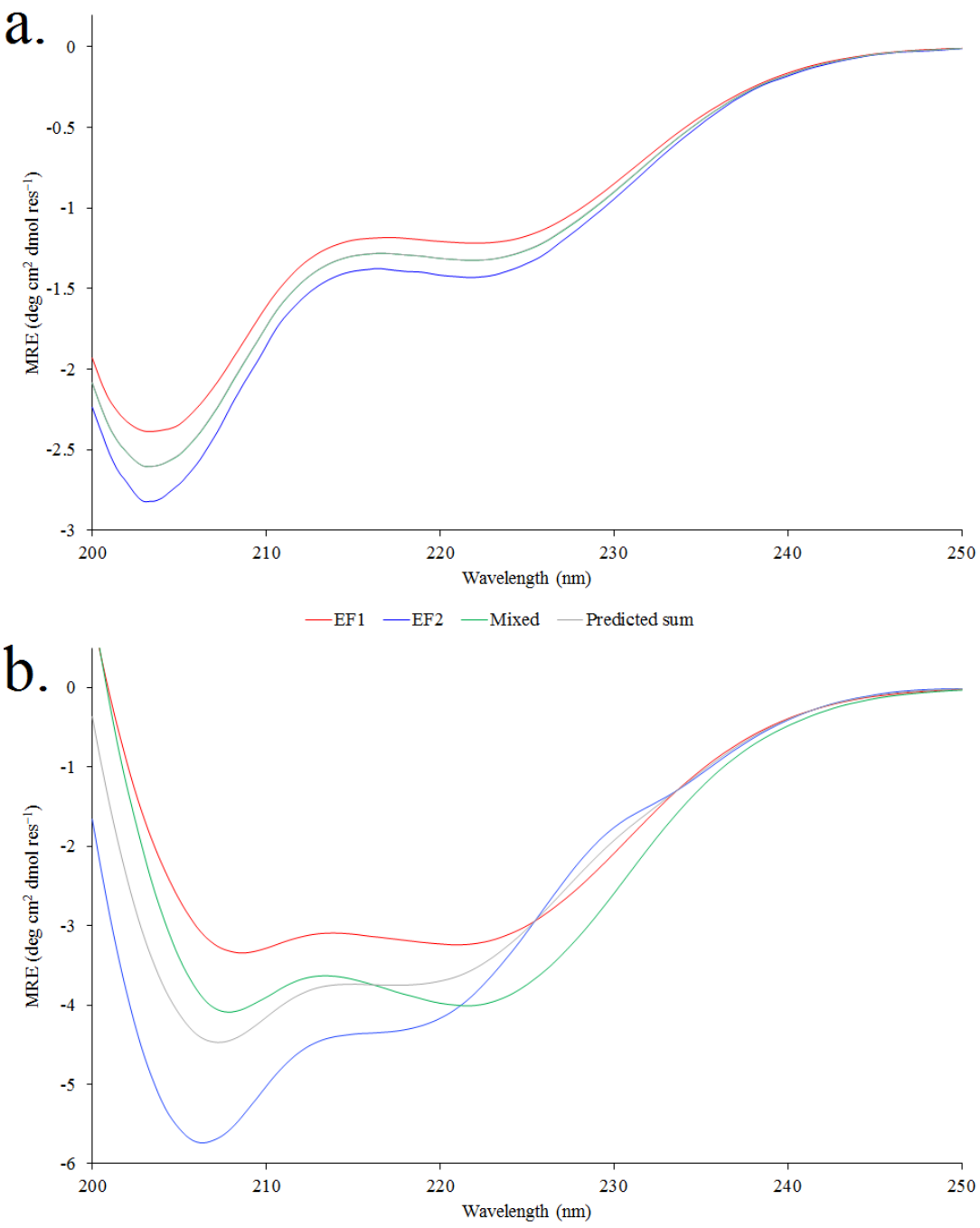


Figure 6.5: CD analysis of Gly-Gly interaction, with either 2 mM $CaCl_2$ or a calcium chelator

a. In the presence of the chelator, the mixed (green spectrum) peptides EF1 (30.4 µM) and S44G/F66W EF2 (25.2 µM) have comparable level of ordered structure to that predicted (grey spectrum) from the spectra of the individual peptides (EF1 = red spectrum, EF2 ΔM43/S44G/F66W = blue spectrum). Buffer: HEPES (10 mM), Tween (0.005 %), EDTA (0.5 mM), pH 7.0. 2 mm cuvette.

b. With 2 mM calcium, the mixed (green spectrum) peptides EF1 (47.4 µM) and S44G/F66W EF2 (41.5 µM) have more ordered structure than that predicted from individual analysis (EF1 = red spectrum, EF2 ΔM43/S44G/F66W = blue spectrum). Buffer: HEPES (10 mM), Tween (0.005 %), $CaCl_2$ (0.5 mM), pH 7.0. 1 mm cuvette.

6.3 Conformationally-assisted chemical ligation and thioester exchange

Despite observing the thioester exchange intermediate during attempts with CI2 in section 5.2 and only limited ligation in the attempt with RNase A peptides in section 4.2, it was hypothesized that the combination of the demonstrated affinity and conformation upon interaction with calcium could encourage conformationally-assisted ligation. As such, under folding conditions and in the presence of thiols, a native amide bond would form at a cysteine-free site. These renewed attempts to demonstrate conformationally-assisted chemical ligation between split protein peptide fragments utilised EF1 as an N-(Me)sulfonamide peptide, and EF2 with a Gly at its *N*-terminus.

6.3.1 Preparation of peptides

EF1 N-(Me)sulfonamide

EF1 N-(Me)sulfonamide was synthesized on pre-loaded Fmoc-Gly-4-sulfamylbutyryl rink amide AM resin (0.26 mmol g^{-1} loading, 0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Upon completion of synthesis the resin was methylated by solvation in DCM and overnight treatment with TMS-diazomethane in hexanes (2.0 M). Cleaved from the resin, resuspended in the denaturing buffer with TCEP-HCl (50 mM) and purified by preparative-HPLC: RP-C18 with a gradient of 15 – 35 % B in 5 min, 35 – 60 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (18.2 mg, 3.13 % yield).

6.3.2 Ligation with $CaCl_2$

The published buffer in which the very low K_d value was obtained contained 2 mM $CaCl_2$. Thus the calcium non-denaturing buffer was employed, where EDTA is replaced with $CaCl_2$ (2 mM).

EF1 N-(Me)sulfonamide (1 mM, 0.48 mg) was added to EF2 ΔM43/S44G/F66W-PEG biotin (1 mM, 0.38 mg) in degassed calcium non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (300 mM), pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.6). The ligation was complete in 8 h, with the product collected and characterized by MALDI-TOF analysis: observed mass = 8844.3 Da in positive linear mode, calculated mass = 8840.96 Da.

6.3.3 Ligation with EDTA

As a control the ligation was repeated in an EDTA-containing buffer (2 mM), the biophysical data suggesting that the presence coordination of calcium will be critical.

EF1 N-(Me)sulfonamide (1 mM, 0.35 mg) was added to EF2 ΔM43/S44G/F66W-PEG biotin (1 mM, 0.27 mg) in degassed non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (300 mM), pH 7.0. The reaction was performed at 30 °C to ensure a constant temperature and monitored by analytical-HPLC: RP-C18 column with a linear gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.7). There is evidence of the ligation product from the 1 h time point, and there appears to be a slight

increase between the 8 h and 24 h time points though the MPAA seems to have oxidised during this time. The ligation product was collected and characterized by MALDI-TOF analysis: 8841.5 Da in positive linear mode, calculated mass = 8840.96 Da.

A comparison of the 4 h time points of ligations with and without calcium shows an increased rate with 2 mM calcium (Figure 6.8).

6.3.4 Ligation in denaturing buffer

As a further control the use of a buffer containing GdnHCl to abolish any structure establishes whether conformation of the peptides is required, even if calcium is present. As such the above experiment was repeated to keep all conditions identical aside from the inclusion of GdnHCl.

EF1N-(Me)sulfonamide (1 mM, 0.22 mg) was added to EF2 ΔM43/S44G/F66W-NH_2 (1 mM, 0.17 mg) in degassed denaturing buffer with TCEP-HCl (50 mM), MPAA (300 mM) and $CaCl_2$ (2 mM), pH 7.5. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 60 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.9). After 24 h there appears to be no obvious change in the abundance of both starting materials.

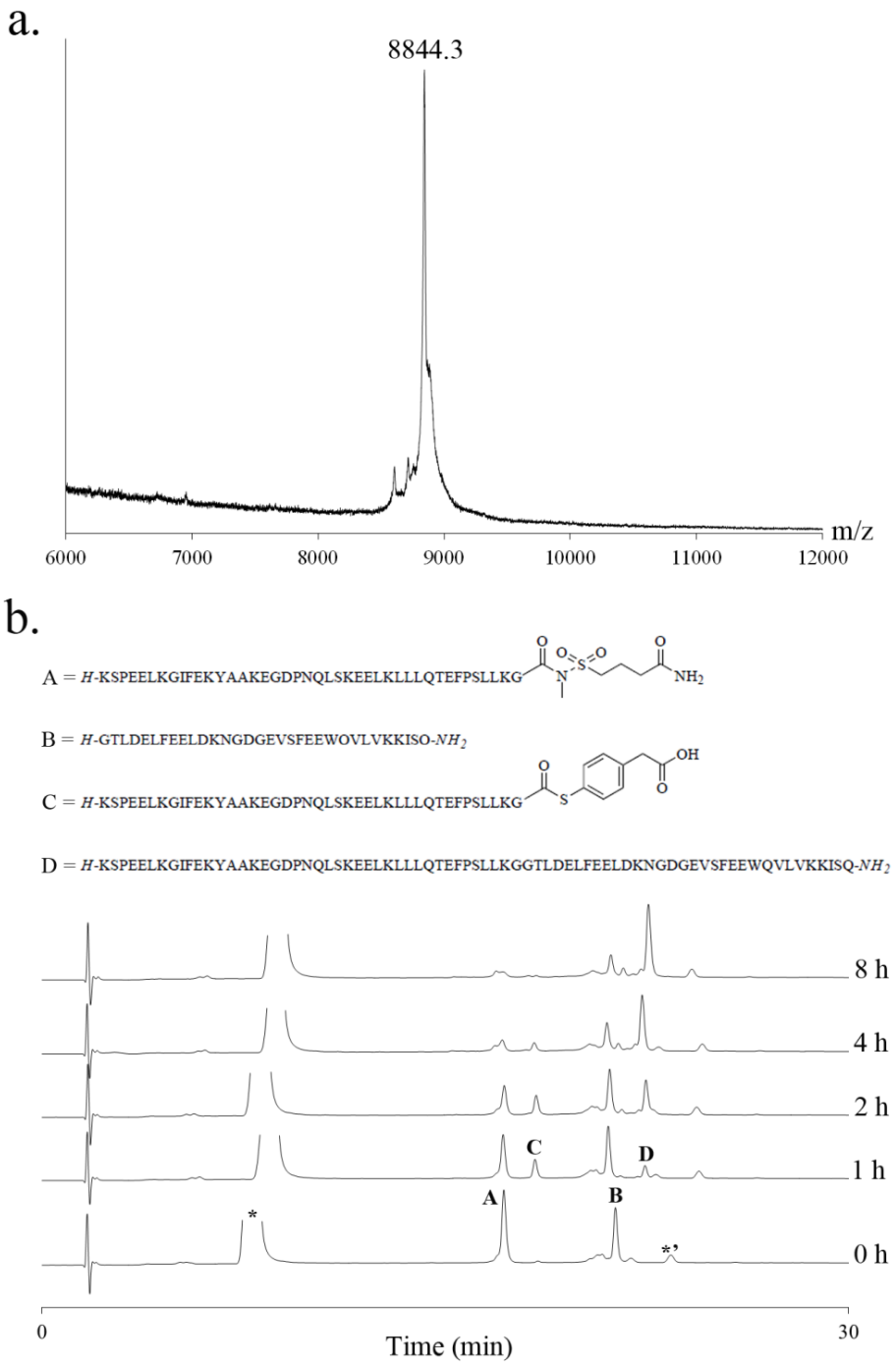


Figure 6.6: Conformationally-assisted ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, with $CaCl_2$

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8844.3 Da, calculated product mass = 8840.96 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (300 mM), $CaCl_2$ (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin, C = EF1 MPAA thioester, D = ligation product, * = MPAA (peak truncated), *' = MPAA impurity.

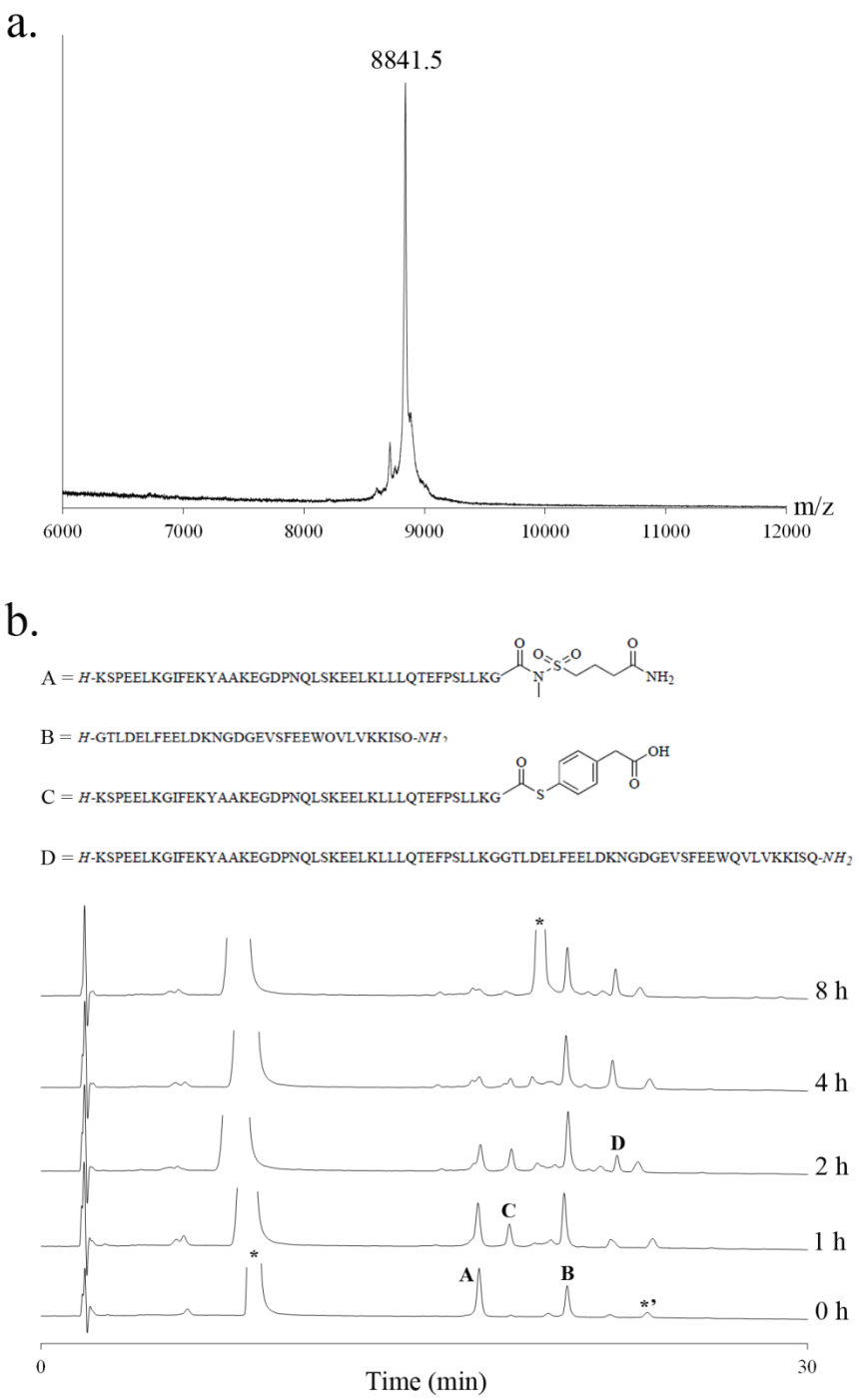


Figure 6.7: Conformationally-assisted ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, with EDTA

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8841.5 Da, calculated product mass = 8840.96 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (300 mM), EDTA (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin, C = EF1 MPAA thioester, D = ligation product, * = MPAA (peak truncated), *' = MPAA impurity.

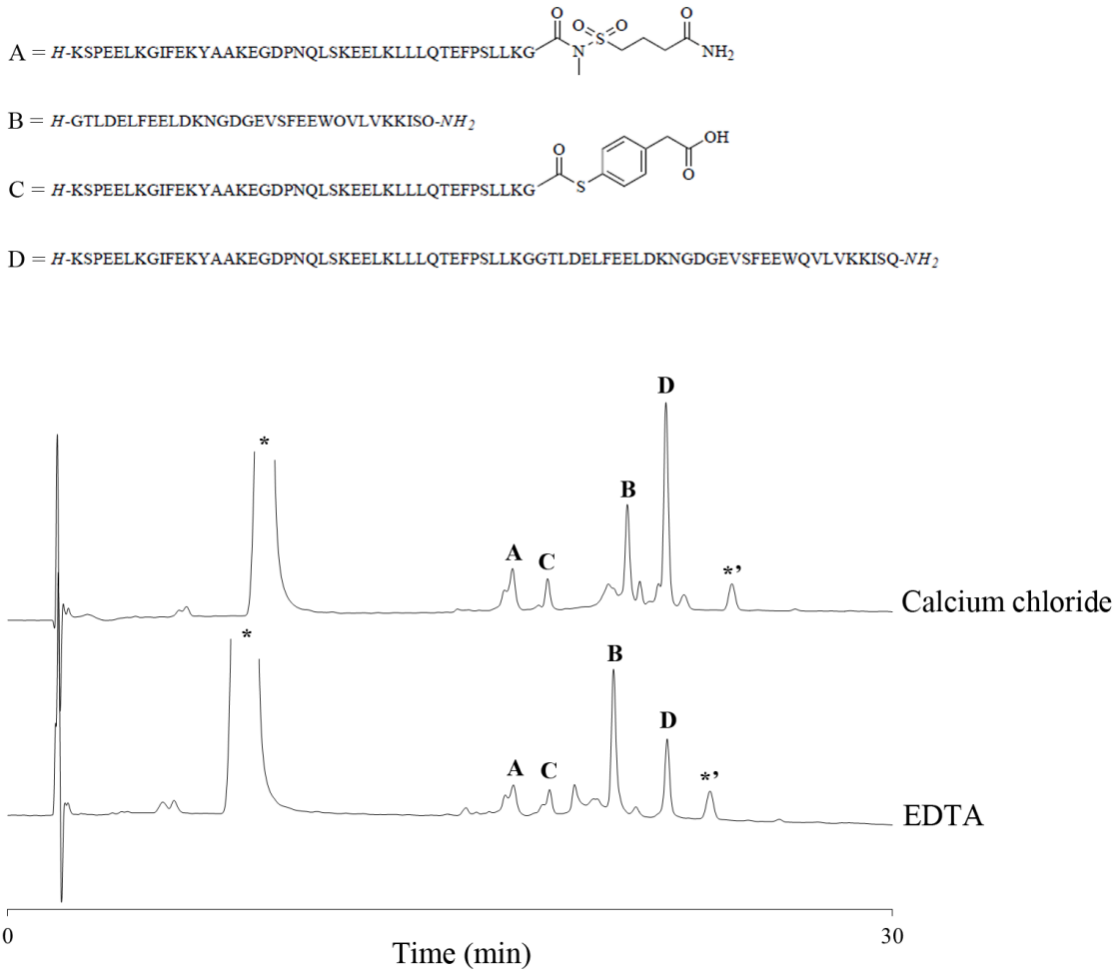


Figure 6.8: Comparison of conformationally-assisted ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, with either $CaCl_2$ or chelator

A comparison of conformationally-assisted ligation under 2 mM calcium (top trace) and 2 mM EDTA (bottom trace). Both traces are the 4 h time point of each reaction. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin, C = EF1 MPAA thioester, D = ligation product, * = MPAA (peak truncated), *' = MPAA impurity.

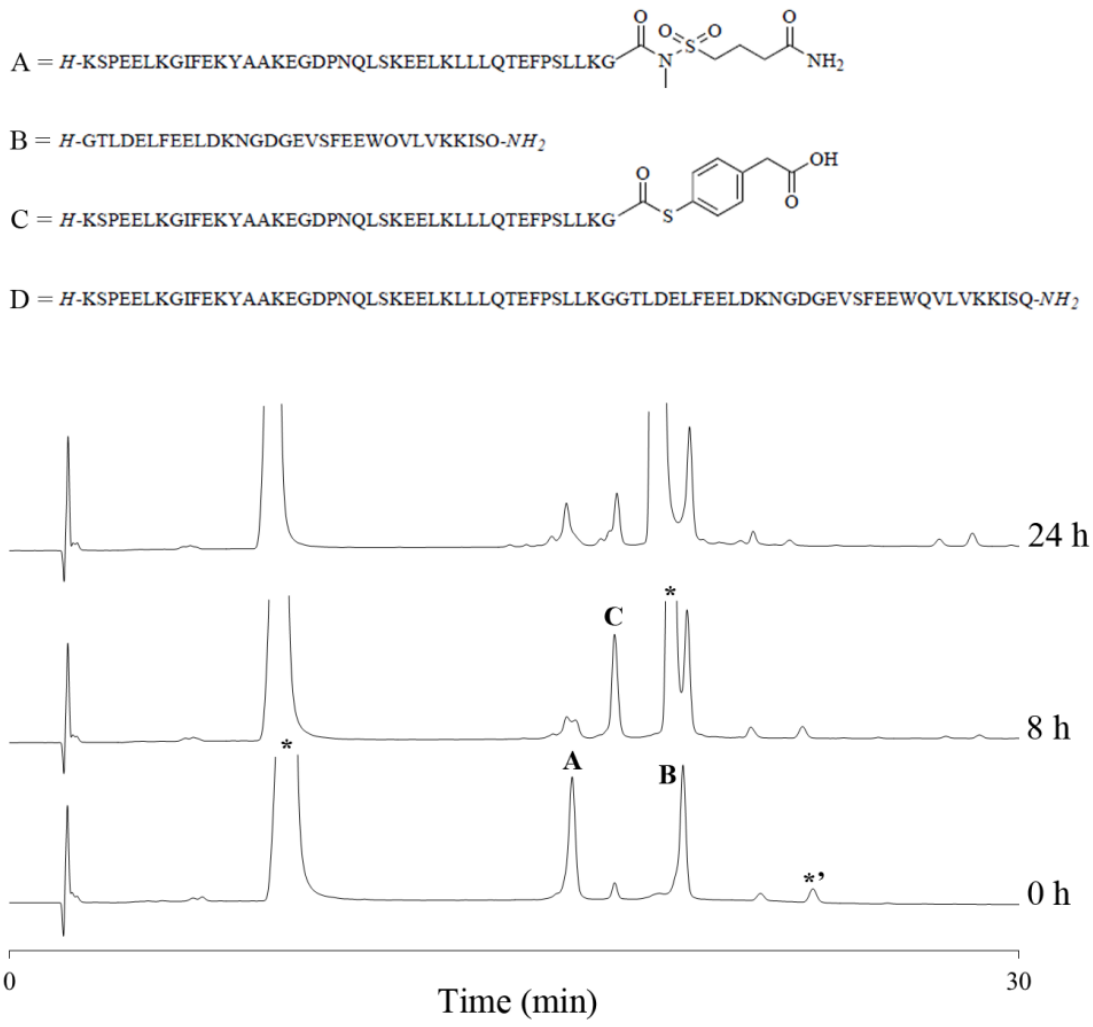


Figure 6.9: Conformationally-assisted peptide ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, in denaturing buffer

Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), GdnHCl (6 M), sodium phosphate (200 mM), TCEP-HCl (50 mM), MPAA (300 mM) and $CaCl_2$ (2 mM), pH 7.5, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin, C = EF1 MPAA thioester, D = ligation product, * = MPAA (peak truncated), *' = MPAA impurity.

6.3.5 Ligation without thiols, with either $CaCl_2$ or a chelator

If conditions can be established whereby the addition of calcium is enough to switch the reaction on, this will provide greater synthetic control. To examine this experiments were set up in parallel whereby the extent of ligation was monitored in the absence of all thiols, with either 2 mM $CaCl_2$ or 2 mM EDTA as a control.

Firstly EF1 N-(Me)sulfonamide (1 mM, 0.34 mg) was added to EF2 ΔM43/S44G/F66W-PEG biotin (1 mM, 0.27 mg) in degassed non-denaturing buffer, pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 column with a linear gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.10 a.). Under these conditions no ligation is seen to take place.

Following this EF1 N-(Me)sulfonamide (1 mM, 0.27 mg) was added to EF2 ΔM43/S44G/F66W-PEG biotin (1 mM, 0.21 mg) in degassed calcium non-denaturing buffer, pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 column with a linear gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.10 b.). Unfortunately this reaction also did not provide ligation product.

6.3.6 Ligation with $CaCl_2$, MESNA

Previous chapters have explored the viability of different thiols in ligation, and so MPAA was replaced with MESNA (10 % w/v) (Adams et al., 2013) as the thiol catalyst.

EF1 N-(Me)sulfonamide (1 mM, 0.26 mg) was added to EF2 ΔM43/S44G/F66W-PEG biotin (1 mM, 0.20 mg) in degassed calcium non-denaturing buffer with TCEP-HCl (50 mM) and MESNA (10 % w/v), pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 column with a linear gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.11). The thioester exchange step is complete after 8 h, but once again there is very little ligation product. The product was collected and confirmed by MALDI-TOF analysis: observed mass = 8844.6 Da in positive linear mode, calculated mass = 8840.96 Da.

6.3.7 Ligation with $CaCl_2$, less MPAA

The concentration of MPAA was lowered to test the effect of peptide structure with a reduced amount of thiol.

EF1 N-(Me)sulfonamide (1 mM, 0.35 mg) was added to EF2 ΔM43/S44G/F66W-PEG biotin (1 mM, 0.28 mg) in degassed calcium non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (60 mM), pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 column with a linear gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.12). Ligation product appears after 2 h and is not complete until around 24 h, and is collected for confirmation by MALDI-TOF analysis: observed mass = 8842.3 Da in positive linear mode, calculated mass = 8840.96 Da.

6.3.8 Ligation with $CaCl_2$, 100 μM peptides

The ligation with calcium non-denaturing buffer and 300 mM MPAA was complete in 8 h with the peptides at 1 mM. The concentration of the peptides was lowered to establish the threshold for ligation under these conditions.

EF1 N-(Me)sulfonamide (100 μM, 0.35 mg) was added to EF2 ΔM43/S44G/F66W-NH_2 (100 μM, 0.25 mg) in degassed calcium non-denaturing buffer with TCEP-HCl (50 mM) and MPAA (300 mM), pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 column with a linear gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1} (Figure 6.13). A problem encountered with this experiment is the amount of sample to load onto the analytical-HPLC: neat overloads the column with MPAA but the peptide peaks are visible, but when diluted the peptide peaks are too small to distinguish above the baseline.

6.3.9 EF1 N-(Me)sulfonamide exchange with MESNA

The thioester exchange intermediate was observed in first the CI2 experiments, and now with the calbindin D_{9k} attempts. Investigating conformationally-assisted chemical ligation with MESNA revealed the thioester exchange step to be complete in 8 h, so these conditions were used to examine this further.

EF1 N-(Me)sulfonamide (1 mM, 4.32 mg) was resuspended in degassed calcium non-denaturing buffer with TCEP-HCl (50 mM) and MESNA (600 mM), pH 7.0. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 μl reaction mixture diluted with 8 μl 0.1 % TFA for injection (Figure 6.14). The exchange product was collected and characterized by MALDI-TOF analysis: observed mass = 4865.9 Da in

positive linear mode, calculated mass = 4871.58 Da. After 5 h the product was purified by preparative-HPLC: RP-C18 column with a gradient of 15-35 % B in 5 min, 35 - 60 % B in 30 min and a flow rate of 15 ml min^{-1} to obtain EF1 MESNA thioester (1.75 mg, 40.6 % yield).

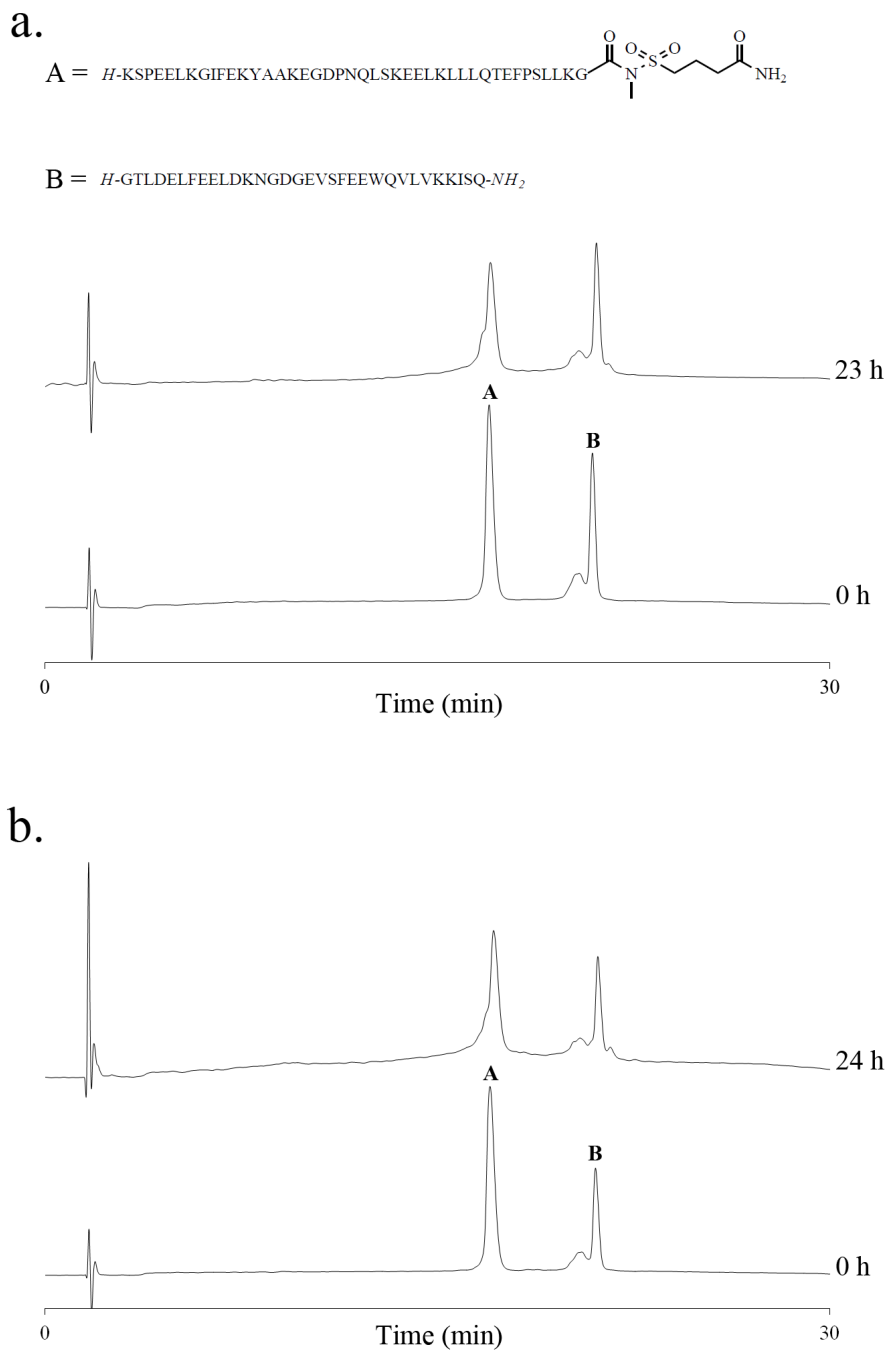


Figure 6.10: Conformationally-assisted peptide ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, without thiols

a. Analytical-HPLC time course of ligation with EDTA. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), EDTA (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin.

b. Analytical-HPLC time course of ligation with calcium. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin.

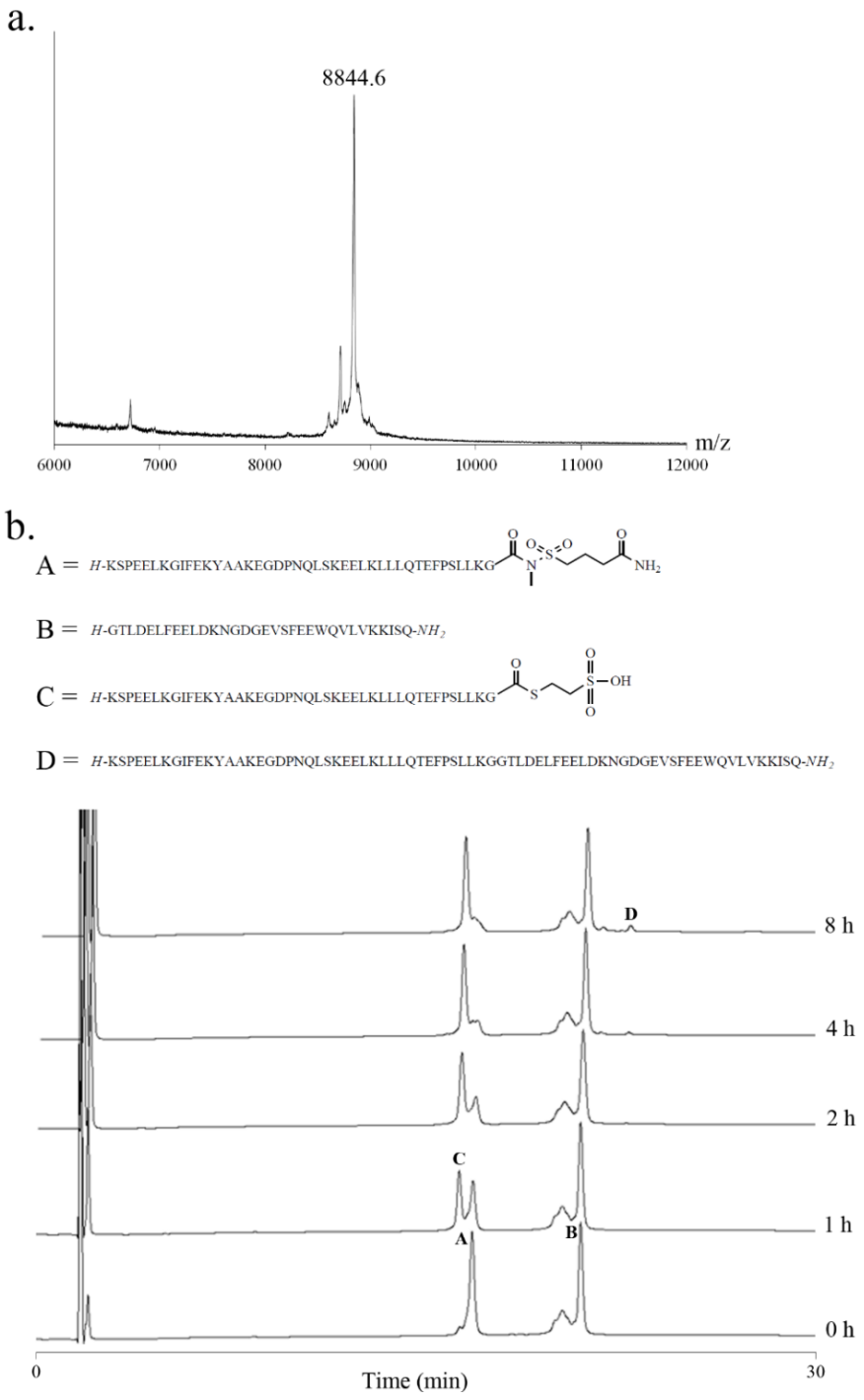


Figure 6.11: Conformationally-assisted ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, with MESNA

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8844.6 Da, calculated product mass = 8840.96 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MESNA (10 % w/v), $CaCl_2$ (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin, C = EF1 MESNA thioester, D = ligation product.

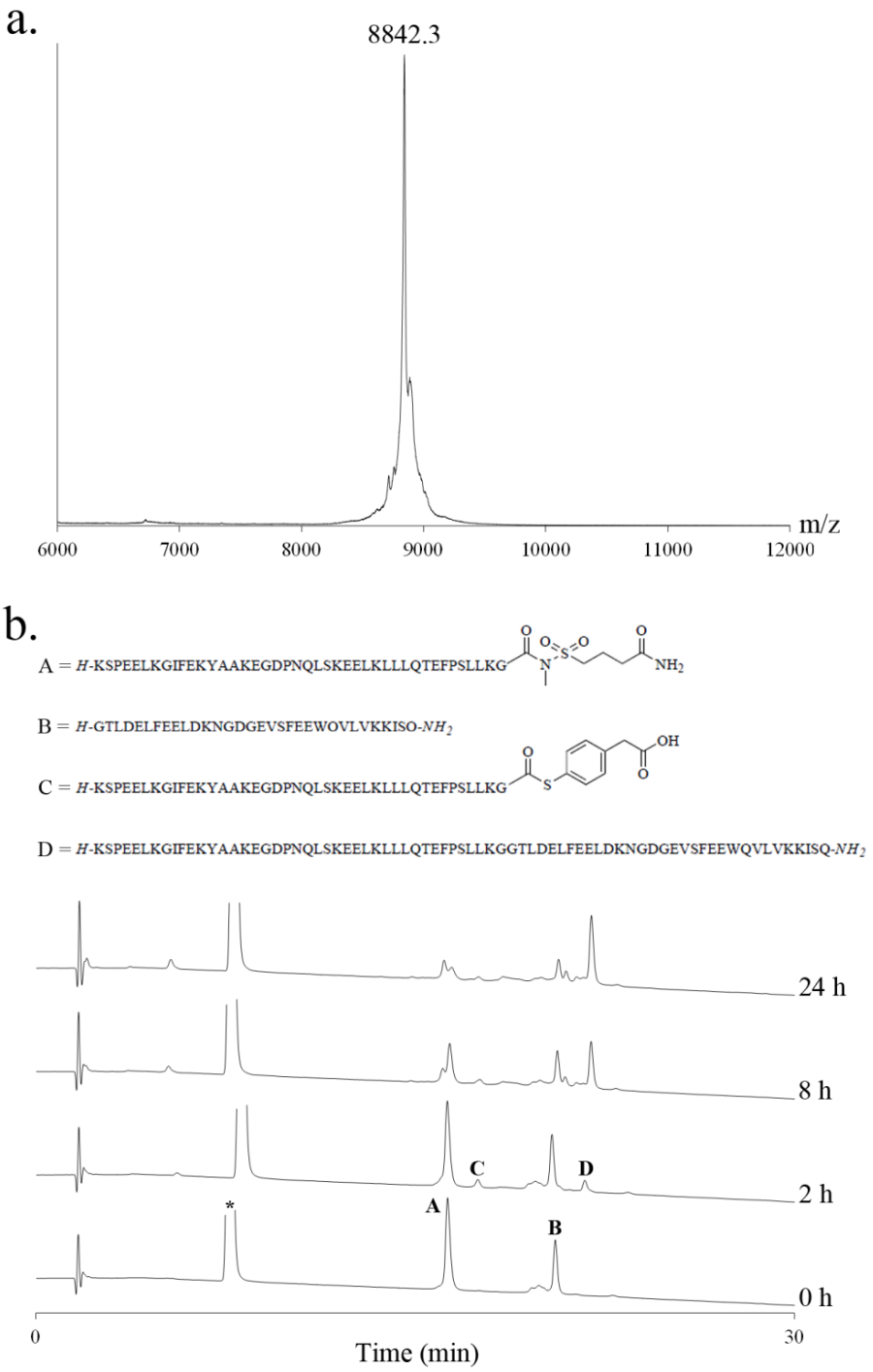


Figure 6.12: Conformationally-assisted ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, with $CaCl_2$, 60 mM MPAA

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8842.3 Da, calculated product mass = 8840.96 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (60 mM), $CaCl_2$ (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF2 ΔM43/S44G/F66W-PEG biotin, C = EF1 MPAA thioester, D = ligation product, * = MPAA (peak truncated).

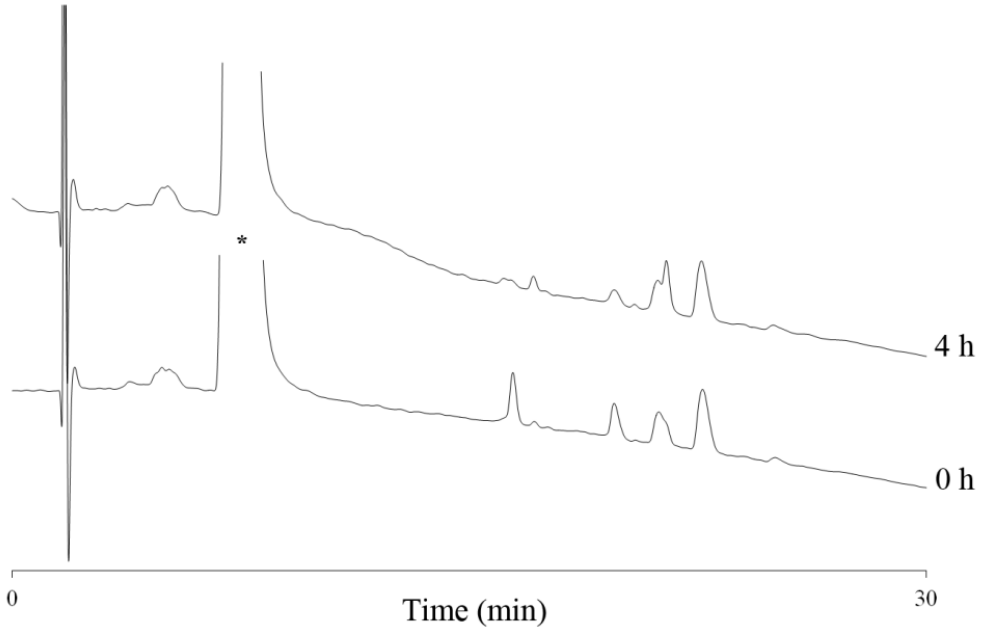


Figure 6.13: Conformationally-assisted ligation of EF1 N-(Me)sulfonamide with EF2 ΔM43/S44G/F66W-PEG biotin, 100 µM peptides

Analytical-HPLC time course of ligation. Ligation conditions: peptides (100 µM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MPAA (300 mM), $CaCl_2$ (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. * = MPAA (peak truncated), other peaks unconfirmed.

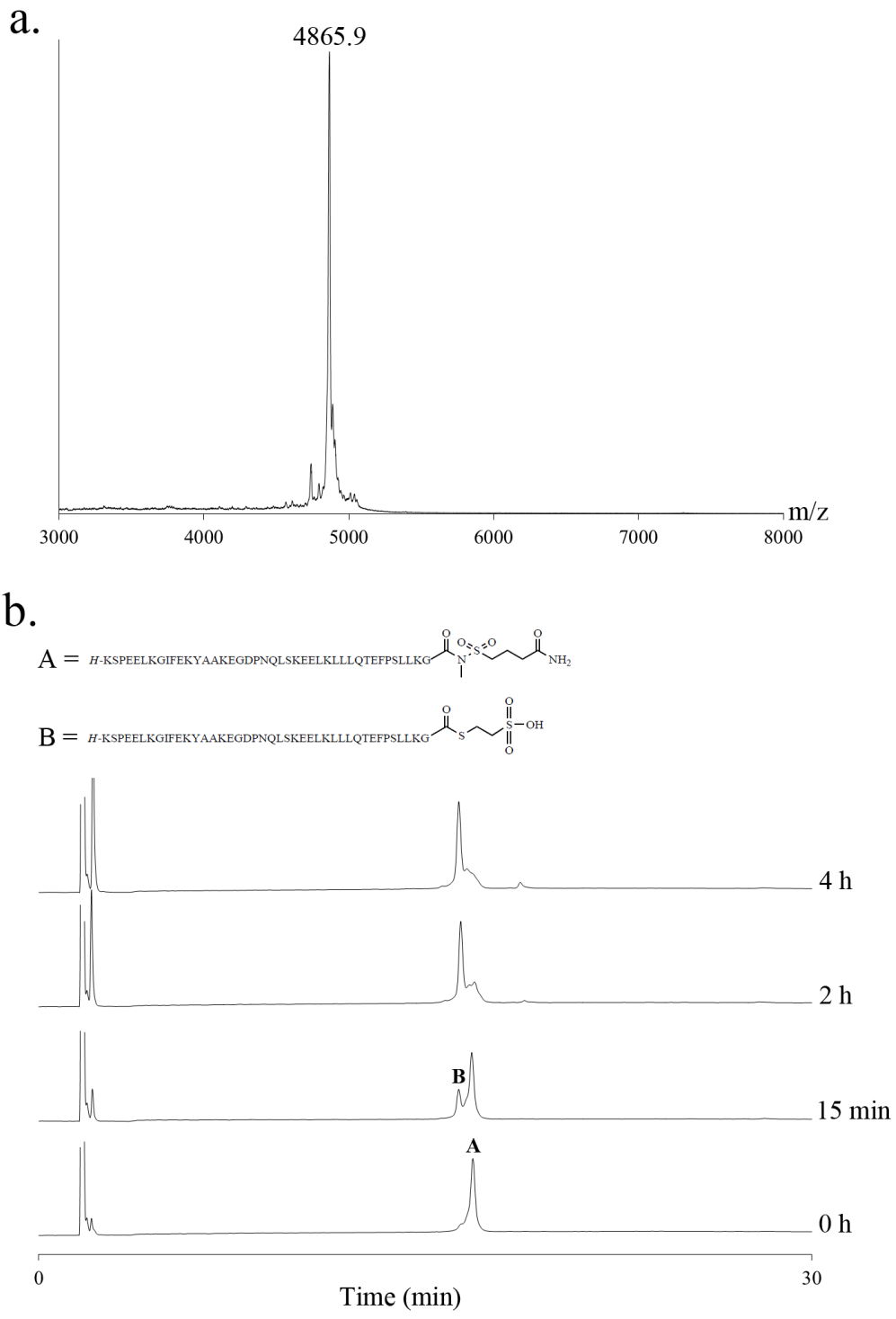


Figure 6.14: EF1 N-(Me)sulfonamide exchange with MESNA

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 4865.9 Da, calculated product mass = 4871.58 Da.

b. Analytical-HPLC time course of exchange. Exchange conditions: peptide (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), MESNA (600 mM), $CaCl_2$ (2 mM), pH 7.0, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF1 MESNA thioester.

6.4 Proximity-assisted acyl transfer

The proposed labelling method would use a proximity-assisted chemical reaction to transfer a tag from a donor to an acceptor peptide bound to the protein of interest. Here we investigate proximity-assisted acyl transfer of a small peptide tag, LAPAG-MPAL thioester, from EF1 to EF2. A modified EF1 with cysteamine at its *C*-terminus is loaded with the tag. Under folding conditions and with 2 mM calcium the conformation of the two calbindin D_{9k} peptides should coordinate their reacting site termini for tag transfer.

6.4.1 EF1 N-(Me)sulfonamide exchange with cysteamine hydrochloride

Cysteamine hydrochloride (20 eq, 2.7 mg) was added to EF1 N-(Me)sulfonamide (1 mM, 7.0 mg) in degassed calcium non-denaturing buffer with TCEP-HCl (50 mM), pH 7.0. The reaction was performed at 25 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.15). Exchange is complete by 7 h, with the product collected for characterization by MALDI-TOF analysis: observed mass = 4803.8 Da in positive linear mode, calculated mass = 4806.53 Da.

6.4.2 Acyl transfer of tag between EF1 cysteamine and EF2 ΔM43/S44G/F66W

EF1 cysteamine (1 mM, 0.53 mg) was mixed with LAPAG-MPAL thioester (1 mM, 0.07 mg) and EF2 ΔM43/S44G/F66W-NH_2 (1 mM, 0.39 mg) in degassed calcium non-denaturing buffer at pH 7.0-7.5. The tag was consumed while some EF1

cysteamine remained, so more LAPAG-MPAL thioester (made up to 10 eq total) was added at 5 h. The reaction was performed at 30 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection (Figure 6.16). There is evidence of acyl transfer from 8 h, with potentially two peaks corresponding to the transfer product. Both were collected for MALDI-TOF analysis: observed mass of peak 1 = 4085.8 Da and observed mass of peak 2 = 4088.1 Da in positive linear mode, calculated mass = 4092.47 Da.

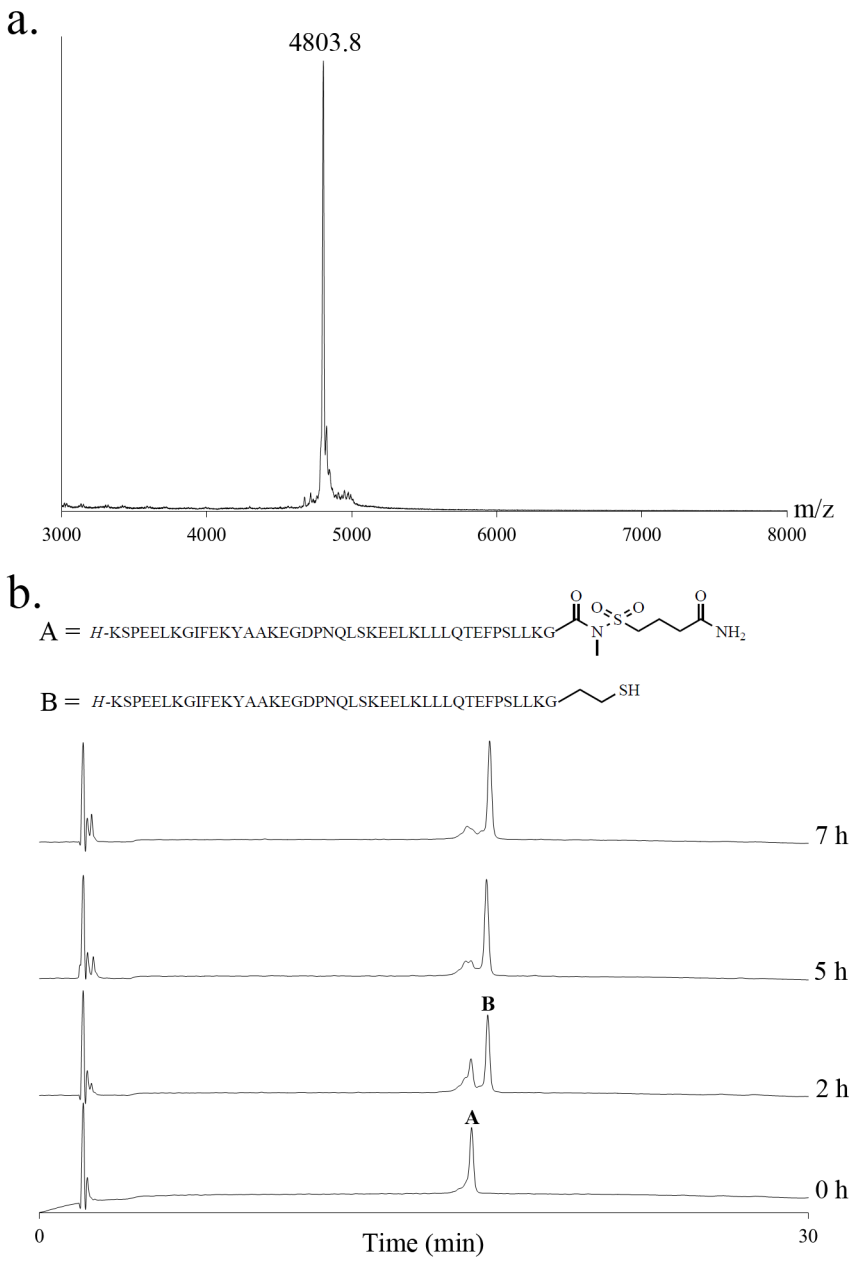


Figure 6.15: EF1 N-(Me)sulfonamide exchange with cysteamine hydrochloride

a. MALDI-TOF MS spectrum of exchange product. m/z identified as 4803.8 Da, calculated product mass = 4806.53 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), TCEP-HCl (50 mM), $CaCl_2$ (2 mM), pH 7.0, 25 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 N-(Me)sulfonamide, B = EF1 cysteamine.

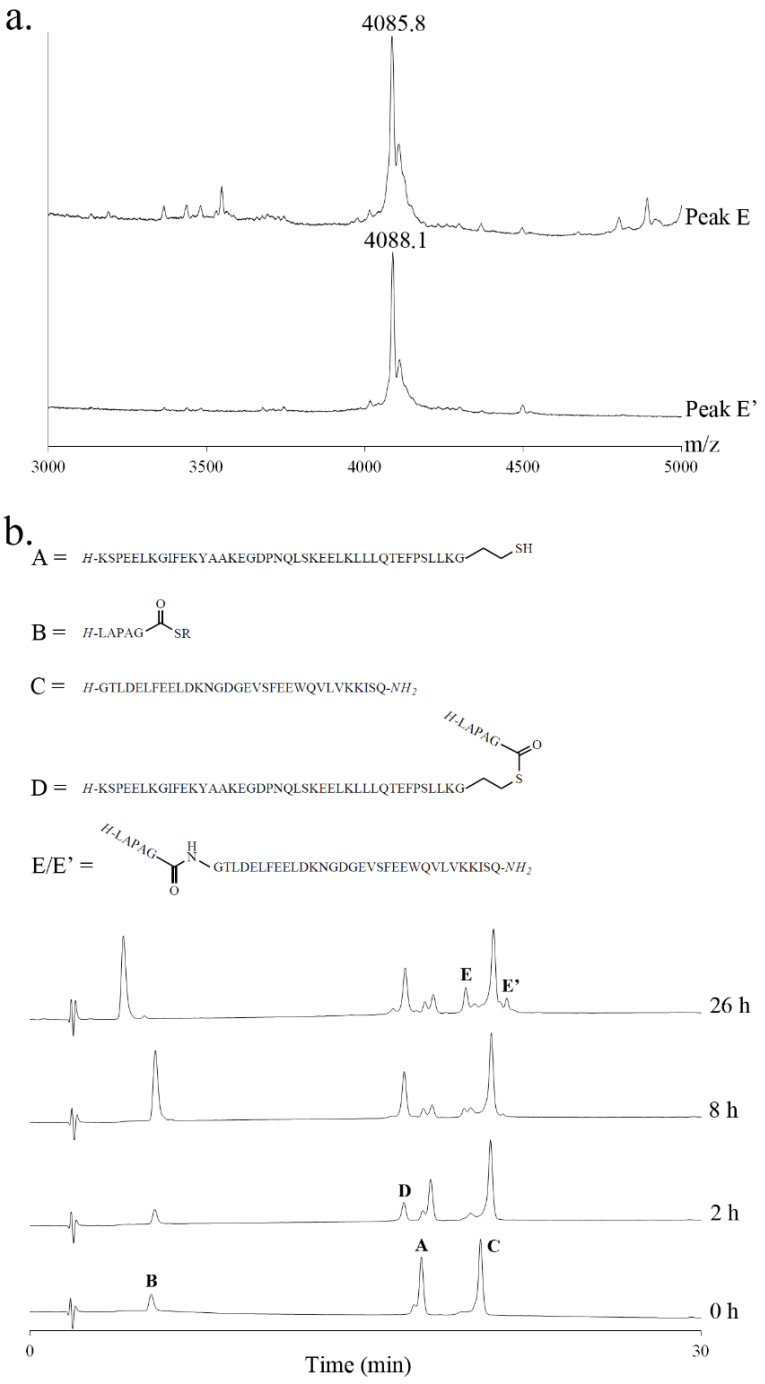


Figure 6.16: Proximity-assisted acyl transfer of tag between EF1 cysteamine and EF2 ΔM43/S44G/F66W-NH_2, with $CaCl_2$

a. MALDI-TOF MS spectrum of acyl transfer product. m/z identified as 4085.8 Da for peak E and 4088.1 Da for peak E', calculated product mass = 4092.47 Da.

b. Analytical-HPLC time course of acyl transfer. Transfer conditions: peptides (1 mM), tag (10 eq) sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.0-7.5, 30 °C. HPLC conditions: RP-C18 column, 20 – 60 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 cysteamine, B = LAPAG thioester, C = EF2 ΔM43/S44G/F66W-NH_2, D = EF1-LAPAG, E/E' = LAPAG-EF2 ΔM43/S44G/F66W-NH_2.

6.5 Conformationally-assisted and Cu-free click chemistry

Two reports of proximity-assisted AAC in the literature are noted in Chapter 1, used in the search for enzyme inhibitors. During the study of inhibitors for HDAC an increased rate was observed and only the 1,4-regioisomer product formed, suggestive of the presence of trace amounts of Cu(I) (Suzuki et al., 2010). Human recombinant HDAC8 was used to further investigate, monitoring by analytical-HPLC. With 1 mM starting materials product was only observed in the presence of HDAC8 unless the active site inhibitor vorinostat was also present, confirming that the components were reacting at the active site. Use of the Cu(I) chelator bathocuprine disulfonic acid inhibited product formation, confirming that Cu(I) is present in the active site and is responsible for the increase in rate observed. The proximity-assisted nature of the reaction was also confirmed by repeating in the presence of Cu but no HDAC8, where no triazole product formed. This was the first example reported whereby a Cu-protein complex induced a proximity-assisted AAC reaction. Earlier work on the search for inhibitors of AChE report a similar proximity-assisted click reaction, though in the absence of Cu (Lewis et al., 2002). The confined binding site of the enzyme acted as a template for 1,3-dipolar cycloaddition, confirmed by the inhibition of product formation when the active site was blocked. The binding site orientates the azide and alkyne components such that a click reaction can take place.

It was hypothesized that the combination of structure and affinity held by the calbindin D_{9k} peptides could encourage conformationally-assisted and Cu-free click ligation. EF1 was synthesized with a *C*-terminal azide and EF2 with an *N*-terminal alkyne.

6.5.1 Preparation of peptides

EF1 azide

EF1 azide was synthesized on Fmoc-azido-NovaTag resin (0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. Cleaved from the resin, resuspended in denaturing buffer and purified by preparative-HPLC: RP-C18 column with a gradient of 15 – 30 % B in 5 min, 30 – 60 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried.

EF2 ΔM43/S44alkyne/F66W

EF2 ΔM43/S44alkyne/F66W was synthesized on PL rink resin (0.39 mmol g^{-1} loading, 0.1 mmol) using microwave-assisted Fmoc-SPPS with DIC/HOBT coupling and piperidine deprotection. In place of the *N*-terminal Gly, propiolic acid was coupled. Addition of propiolic acid (1 mmol, 70 mg) in minimal DMF, DIC in DMSO (0.5 ml of 0.8 M) and HOBT in DMF (1 ml of 0.5 M) resulted in a peptide with *N*-terminal alkyne after 45 min. Cleaved from the resin, resuspended in the denaturing buffer and purified by preparative-HPLC: RP-C18 with a gradient of 15 – 35 % B in 5 min, 35 – 65 % B in 30 min and a flow rate of 15 ml min^{-1}. Purified peptide freeze-dried (6.3 mg, 1.71 % yield).

6.5.2 Ligation with 10 mM peptides, 50 °C

The first attempt at a click reaction was carried out under folding conditions. The peptides were at 1 mM and the temperature at 30 °C, a temperature significantly lower than those observed in traditional click reactions. The reaction was monitored by

analytical-HPLC (RP-C18 with a gradient of 20 – 60 % B in 30 min and a flow rate of 1 ml min^{-1}) but no change was observed, and the ligation appeared unsuccessful (data not shown). To improve the likelihood of successful ligation the peptide concentrations were increased, along with temperature.

EF1 azide (10 mM, 1.7 mg) was added to EF2 ΔM43/S44alkyne/F66W (10 mM, 1.2 mg) in degassed calcium non-denaturing buffer, pH 7.5 and without any additives. The reaction was performed at 50 °C and monitored by analytical-HPLC: RP-C18 with an initial gradient of 20 – 60 % B in 30 min and then a gradient of 20 – 60 % B in 30min, both with a flow rate of 1 ml min^{-1} (Figure 6.17). The gradient change was implemented following the observation of a reduced EF2 peak after 2 h, and so it was lengthened in case the ligation product eluted at a higher percentage of AcN. With this new gradient we can now see product for the 4 h time point onwards, collected for confirmation by MALDI-TOF analysis: observed mass = 8495.5 Da in positive linear mode, calculated mass = 8492.47 Da.

6.5.3 Ligation with 1 mM peptides, 37 °C

Both the concentration of the starting materials and the temperature were now lowered, after observing ligation product with 10 mM peptides and at 50 °C.

EF1 azide (1 mM, 0.4 mg) was added to EF2 ΔM43/S44alkyne/F66W (1 mM, 0.29 mg) in the calcium non-denaturing buffer at pH 7.5, without any additives. In this instance the buffer was not degassed prior to its addition to the peptides. The reaction was performed at 37 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 90 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.18). There is evidence of ligation product in the 1 h trace, and significant amounts of product in a 25 h time point.

The ligation product was collected and characterized by MALDI-TOF analysis: observed mass = 8491.8 Da in positive linear mode, calculated mass = 8492.47 Da.

6.5.4 Ligation with 1 mM peptides, denaturing buffer

As a control experiment similar to that conducted in section 6.3.4 the denaturing buffer was used to remove peptide secondary structure. Even with 2 mM calcium this should confirm the role of peptide structure in this ligation.

EF1 azide (1 mM, 0.35 mg) was added to EF2 ΔM43/S44alkyne/F66W (1 mM, 0.25 mg) in the denaturing buffer, with $CaCl_2$ (2 mM) at pH 7.5. In this instance the buffer was not degassed prior to its addition to the peptides. The reaction was performed at 37 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 90 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.19). After 24 h there appears to be no change in the abundance of either starting material and no ligation product is observed.

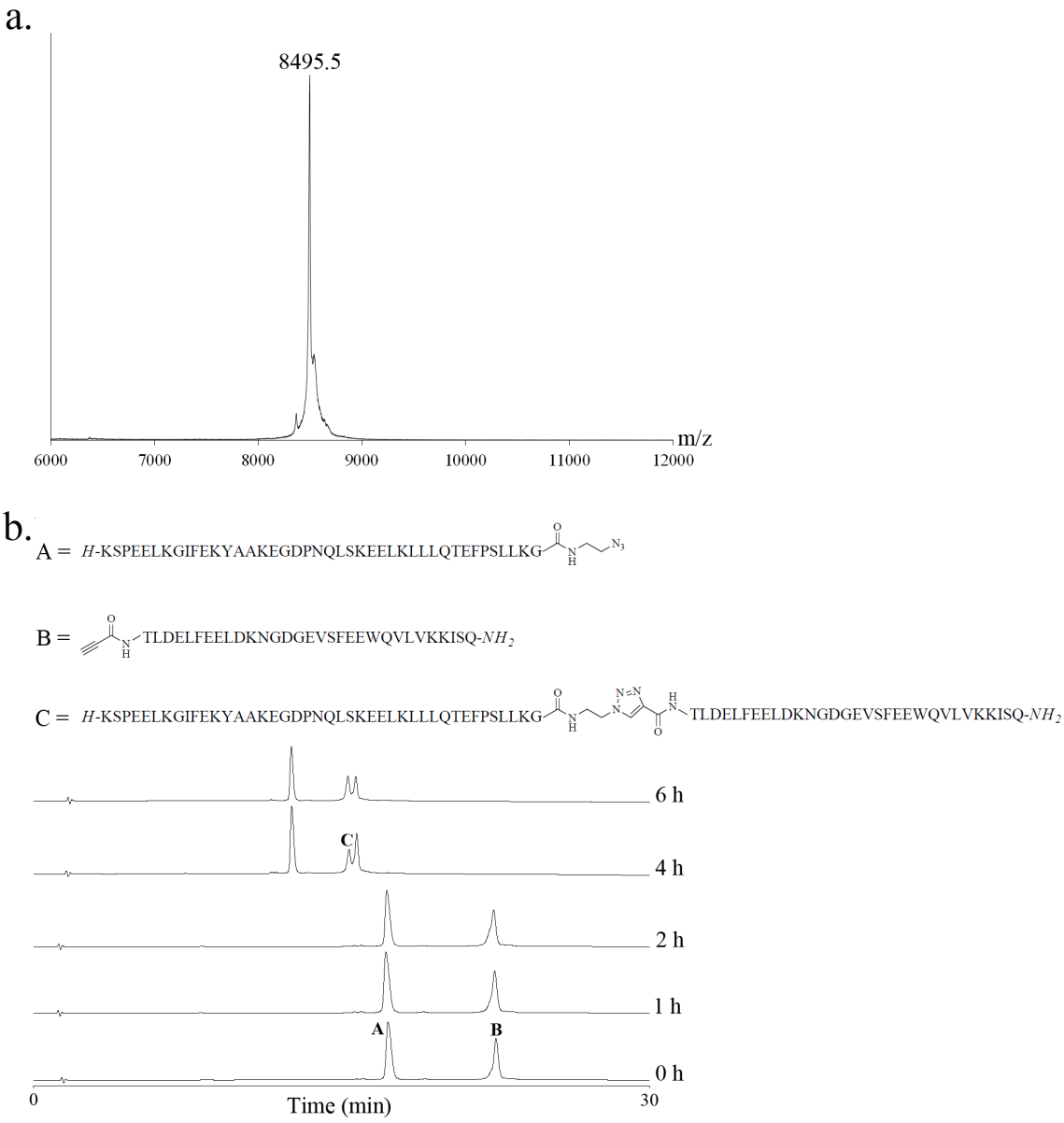


Figure 6.17: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, with 10 mM peptides, 50 °C

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8495.5 Da, calculated product mass = 8492.47 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (0.25 mM EF1 azide, 0.5 mM EF2 ΔM43/S44alkyne/F66W), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.5, 50 °C HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W, C = ligation product.

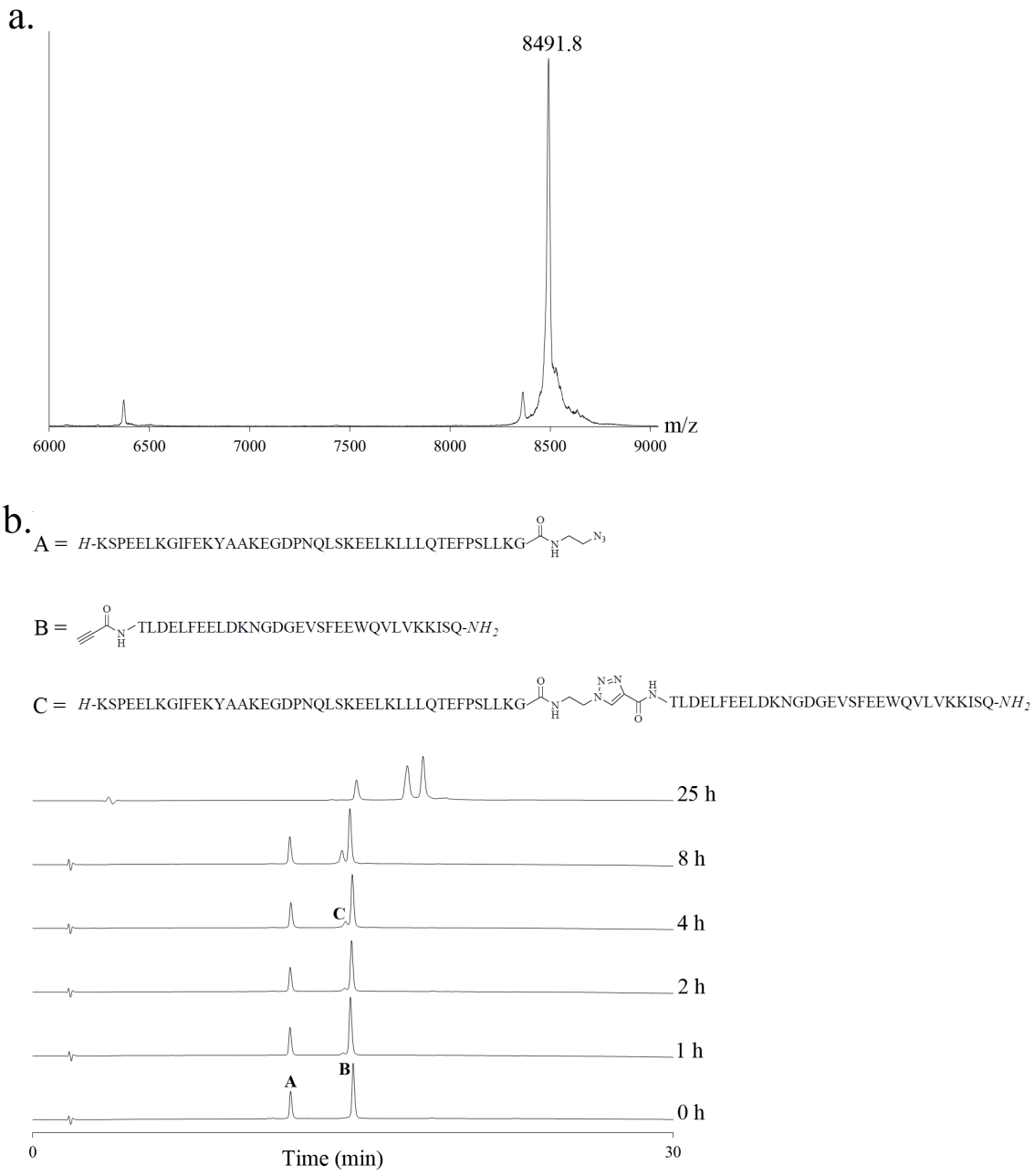


Figure 6.18: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, with 1 mM peptides, 37 °C

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8491.8 Da, calculated product mass = 8492.47 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.5, 37 °C HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W, C = ligation product.

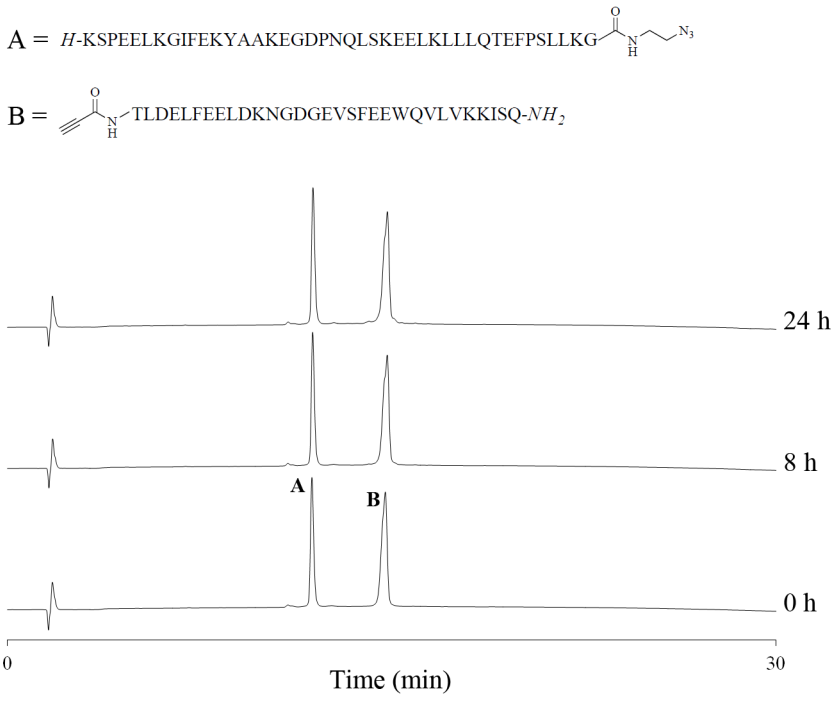


Figure 6.19: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, in denaturing buffer

Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), GdnHCl (6 M), sodium phosphate (200 mM), $CaCl_2$ (2 mM), pH 7.5, 37 °C. HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W.

6.5.5 Ligation with 1 mM peptides, 50 °C

Attempting to increase the rate of the ligation, the temperature was increased to the original 50 °C. As these conditions had previously shown promising progress it was decided to maintain the peptide concentration at 1 mM.

EF1 azide (1 mM, 0.48 mg) was added to EF2 ΔM43/S44alkyne/F66W (1 mM, 0.35 mg) in the calcium non-denaturing buffer at pH 7.5, without any additives or degassing since no negative effects were observed in section 6.5.3. The reaction was

performed at 50 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 90 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.20). Under these conditions the ligation appears complete by 24 h.

6.5.6 Ligation with 0.5 mM peptides, 50 °C

The previous ligation at 50 °C was the best so far, and to investigate the effect of peptide concentration on the rate the concentration of both starting materials was lowered.

EF1 azide (0.5 mM, 0.18 mg) was added to EF2 ΔM43/S44alkyne/F66W (0.5 mM, 0.13 mg) in the calcium non-denaturing buffer at pH 7.5, without any additives or degassing. The reaction was performed at 50 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 90 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.21). Under these conditions the ligation also appears complete by 24 h, and the product collected for confirmation by MALDI-TOF analysis: observed mass = 8494.0 Da in positive linear mode, calculated mass = 8492.47 Da.

6.5.7 Ligation with 0.25 mM EF1 azide, 50 °C

Each ligation attempt appears to have an amount of EF1 azide remaining once EF2 ΔM43/S44alkyne/F66W is depleted, thus its concentration was halved relative to its partner alkyne peptide.

EF1 azide (0.25 mM, 0.17 mg) was added to EF2 ΔM43/S44alkyne/F66W (0.5 mM, 0.24 mg) in the calcium non-denaturing buffer at pH 7.5, without any additives or degassing. The reaction was performed at 50 °C and monitored by analytical-HPLC:

RP-C18 with a gradient of 20 – 90 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.22). Despite reducing the concentration of one of the starting materials these conditions still encourage the completion of ligation by 24 h, and the ligation product was collected for confirmation by MALDI-TOF: observed mass = 8494.2 Da in positive linear mode, calculated mass = 8492.47 Da.

6.5.8 Ligation with 0.25 mM EF1 azide, 45 °C

Having established that the ligation does not appear concentration-dependent at this level the temperature is now lowered to 45 °C.

EF1 azide (0.25 mM, 0.13 mg) was added to EF2 ΔM43/S44alkyne/F66W (0.5 mM, 0.19 mg) in the calcium non-denaturing buffer at pH 7.5, without any additives or degassing. The reaction was performed at 45 °C and monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 90 % B in 30min and a flow rate of 1 ml min^{-1} (Figure 6.23). Unlike the equivalent experiment at 50 °C there is still a substantial amount of both starting materials at 24 h, and the reaction is completed by 48 h. The ligation product was collected and characterized by MALDI-TOF analysis: observed mass = 8494.2 Da in positive linear mode, calculated mass = 8492.47 Da.

Reproducibility was tested by repeating this experiment, and nanodrop was used to confirm the peptide concentrations.

EF1 azide (0.34 mM) and EF2 ΔM43/S44alkyne/F66W (0.65 mM) were each dissolved in half the total volume of the calcium non-denaturing buffer at pH 7.5. Their concentrations were confirmed by nanodrop analysis, and the peptides were then mixed. The reaction was monitored by analytical-HPLC: RP-C18 with a gradient of 20 – 90 %

B in 30 min and a flow rate of 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. The ligation was again complete by 48 h.

This ligation was repeated on a larger scale and the product purified by preparative-HPLC. EF1-azide (0.25 mM, 1.02 mg) and EF2 ΔM43/S44alkyne/F66W (0.5 mM, 1.47 mg) were mixed under the same conditions as section 6.4.2, the reaction monitored at t = 0 and 48 h by analytical-HPLC: RP-C18 20 – 90 % B in 30 min, 1 ml min^{-1}. 2 µl reaction mixture diluted with 8 µl 0.1 % TFA for injection. Product purification by preparative-HPLC: RP-C18 with a gradient of 20 – 90 % B in 30 min, 15 ml min^{-1}. Purified peptide freeze-dried (0.6 mg, 24.1 % yield).

6.5.9 Nuclear magnetic resonance (NMR) spectroscopy

The work in this section was carried out by Dr Geoff Kelly. Following the conformationally-assisted Cu-free click reaction with the azide and alkyne peptides we sought to investigate the distribution of the product triazole, using the purified product described in section 6.5.8. The triazole isomers can be identified by $^{1}H/^{13}C$-HSQC using gated decoupling, with the C_5 signal of the 1,4-regioisomer at δ~120 ppm and the C_4 signal of the 1,5-regioisomer at δ~133 ppm (Creary et al., 2012) (Figure 6.24). The product peptide was prepared as unbuffered sample in H_2O (95 %) and D_2O (5 %), at pH 6.0.

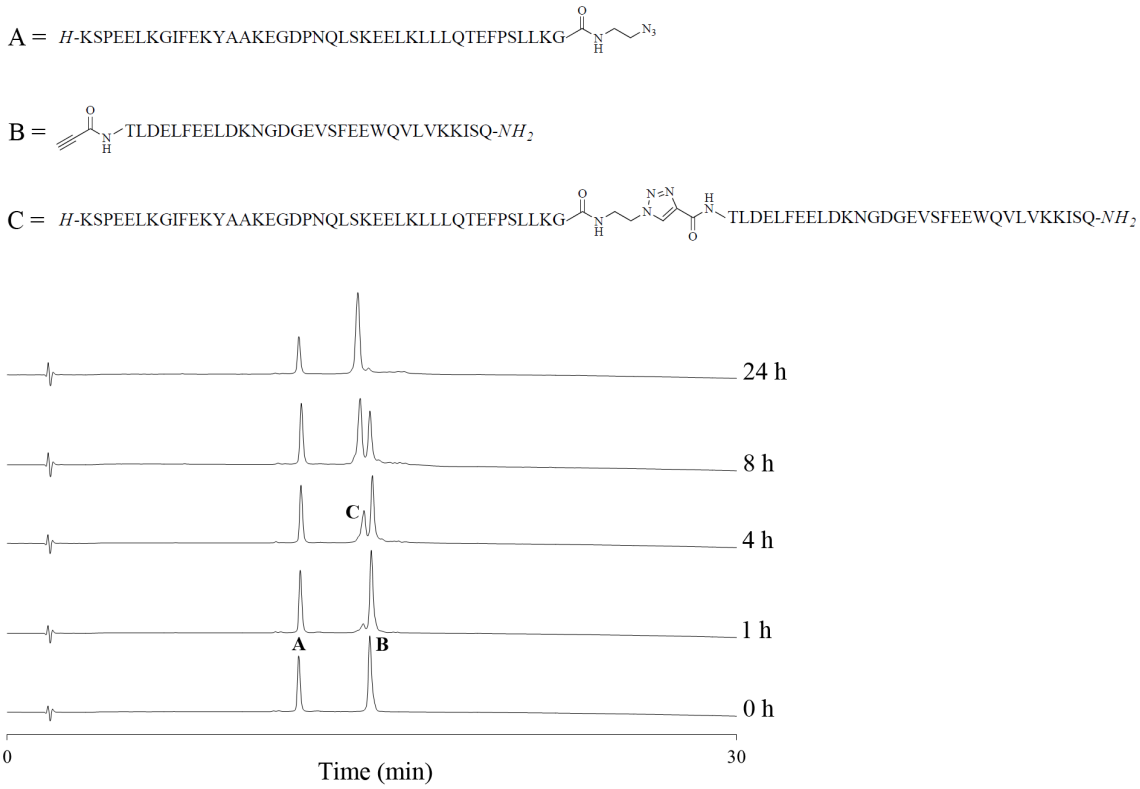


Figure 6.20: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, with 1 mM peptides, 50 °C

Analytical-HPLC time course of ligation. Ligation conditions: peptides (1 mM), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.5, 50 °C HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W, C = ligation product.

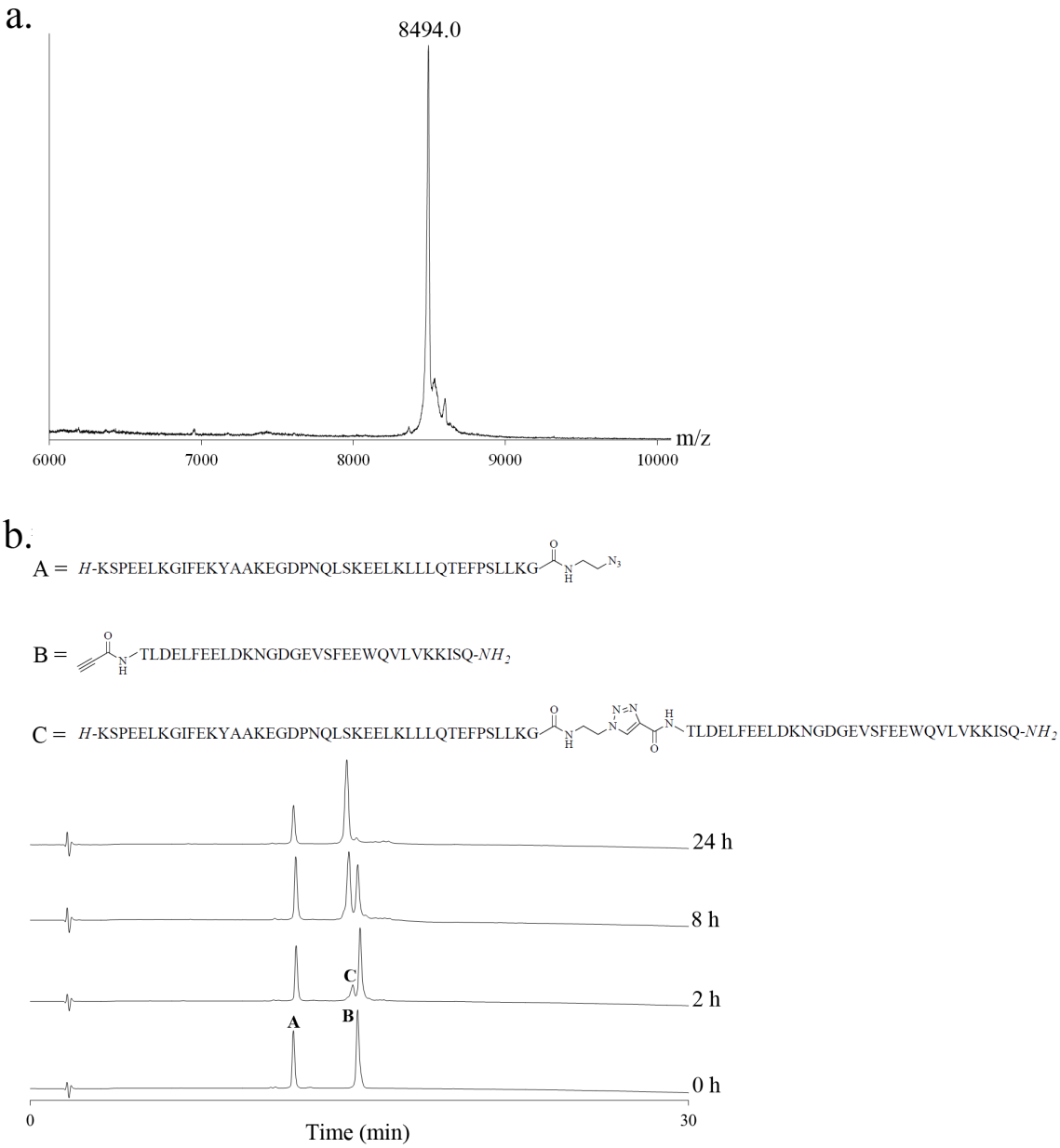


Figure 6.21: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, with 0.5 mM peptides, 50 °C

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8494.0 Da, calculated product mass = 8492.47 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (0.5 mM), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.5, 50 °C HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W, C = ligation product.

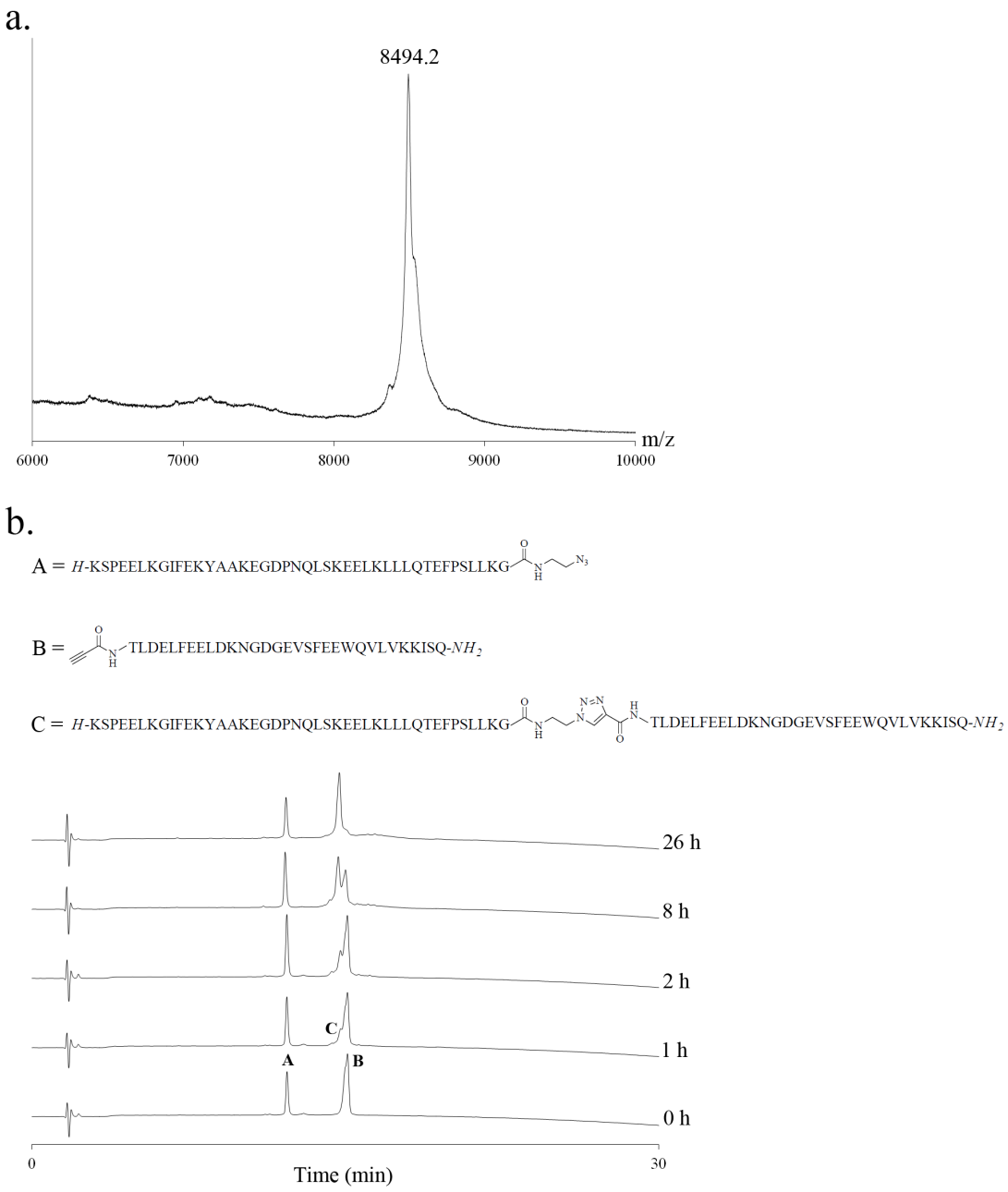


Figure 6.22: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, with 0.25 mM EF1 azide, 50 °C

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8494.2 Da, calculated product mass = 8492.47 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (0.25 mM EF1 azide, 0.5 mM EF2 ΔM43/S44alkyne/F66W), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.5, 50 °C. HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W, C = ligation product.

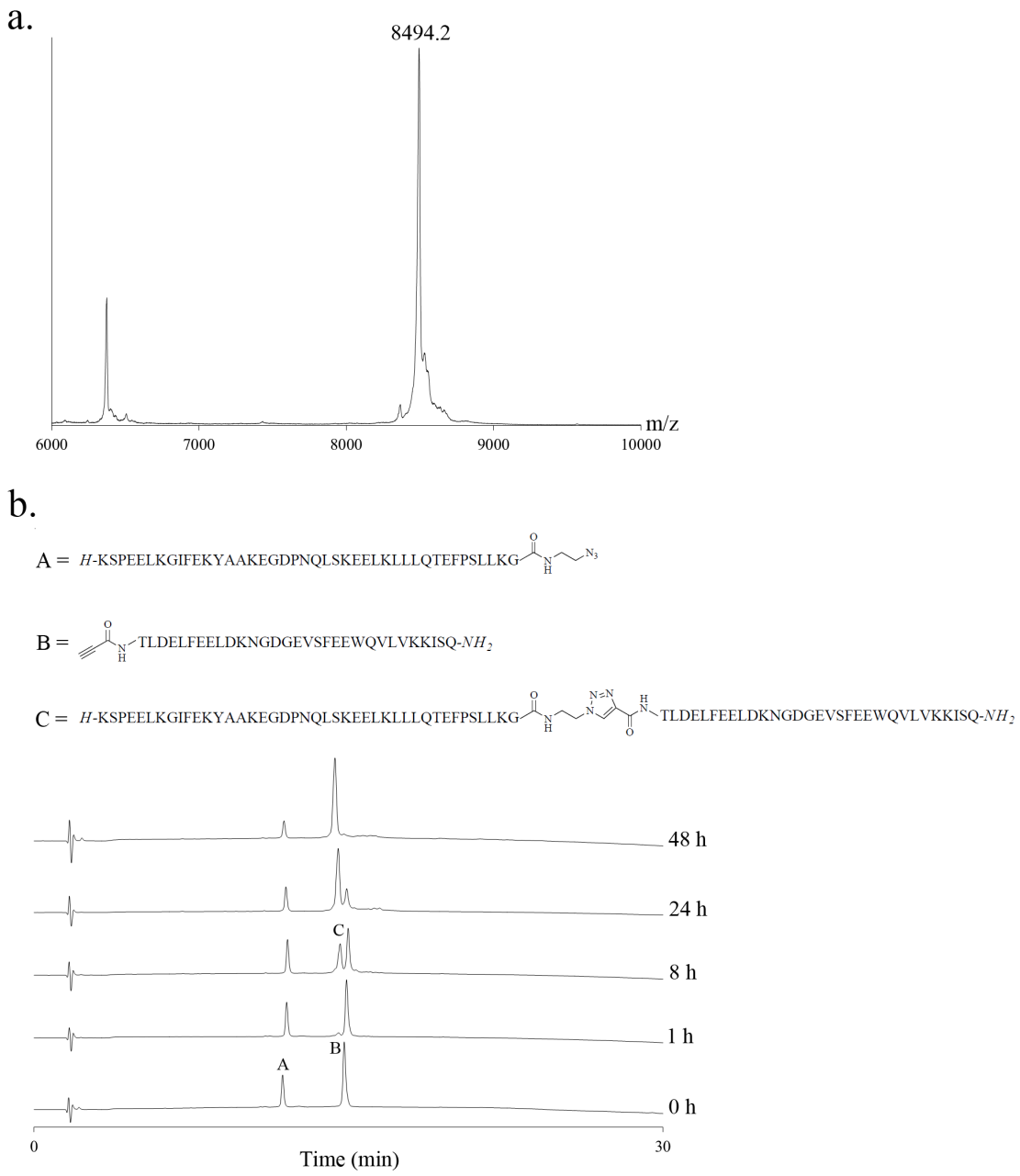


Figure 6.23: Conformationally-assisted and Cu-free click chemistry with EF1 azide and EF2 ΔM43/S44alkyne/F66W, with 0.25 mM EF1 azide, 45 °C

a. MALDI-TOF MS spectrum of ligation product. m/z identified as 8494.2 Da, calculated product mass = 8492.47 Da.

b. Analytical-HPLC time course of ligation. Ligation conditions: peptides (0.25 mM EF1 azide, 0.5 mM EF2 ΔM43/S44alkyne/F66W), sodium phosphate (100 mM), $CaCl_2$ (2 mM), pH 7.5, 45 °C. HPLC conditions: RP-C18 column, 20 – 90 % B (TFA (0.1 %), AcN (90 %)) in 30 min, 1 ml min^{-1}. A = EF1 azide, B = EF2 ΔM43/S44alkyne/F66W, C = ligation product.

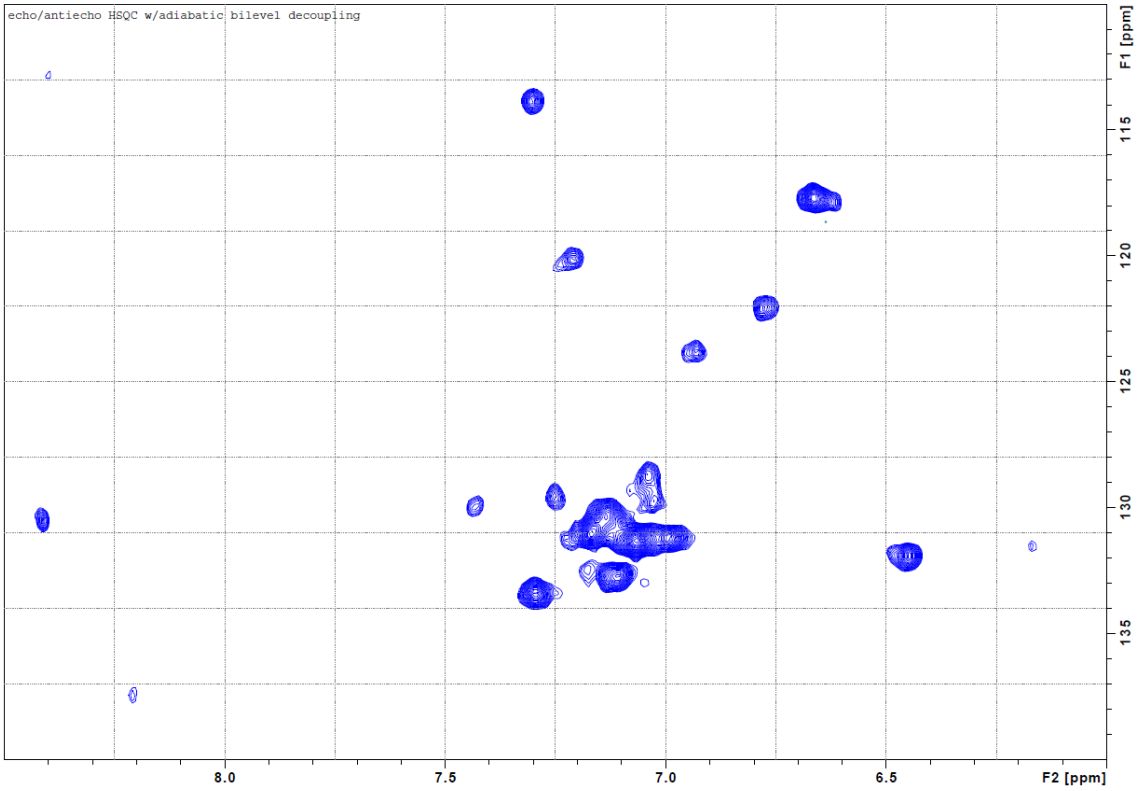


Figure 6.24: $^{1}H/^{13}C$-HSQC NMR spectrum of calbindin D_{9k} Cu-free click product

The aromatic region of a $^{1}H/^{13}C$-HSQC NMR spectrum for the product of the calbindin D_{9k} Cu-free click reaction, displaying signals of protonated carbons. With at least 7 signals in the ~120 ppm region this indicates a 1,4-disubstituted triazole, accounting for signals from Tyr and Trp. To accurately assign signals signal-to-noise must be improved, particularly for those in the ~130 ppm region. Peptide prepared as unbuffered sample in H_2O (95 %) and D_2O (5 %) at pH 6.0.

6.6 Calbindin D_{9k} peptides can perform conformationally-assisted and Cu-free click ligations

6.6.1 Biophysical characterization

Work with the calbindin D_{9k} EF1 and EF2 peptides began by establishing the affinity of modified EF-hand peptides for one another using BLI. As discussed in Chapter 4 BLI is able to provide K_d values in a similar manner to SPR, the technique used in the work discussed in this chapter's introduction. In the paper by Linse and co-workers discussed in the introduction to this chapter, SPR was conducted using a BIAcore biosensor system with a flow buffer of HEPES/KOH (10 mM, pH 7.4), NaCl (0.15 M), Tween 20 (0.005 %), NaN_3 (0.02 %) and either $CaCl_2$ (2 mM) or EDTA (3.4 mM). Previous BLI experiments with the split proteins in earlier chapters immobilized the *N*-terminal fragment, however here it was decided to immobilize EF2 rather than EF1. Linse and colleagues measured the interaction between EF1 and immobilized EF2 for different EF1 variants. They also monitored the interaction between EF2 and immobilized EF1. The dissociation rate varied between mutants and was not concentration-dependent, but the association rate was also not found to be concentration dependent. Therefore association is governed by a separate rate-limiting step, thought to be the dissociation of EF1 dimers.

The first experiment, Gly-Met, was chosen based upon the P43M mutation used to separate the calbindin D_{9k} peptides (Finn et al., 1992). For this interaction site a K_d value of 47.7 nM ± 5.8 nM was determined (Figure 6.1), which is significantly higher than that published. We wondered if the cleavage by CNBr at this mutated P43M site resulted in a homoserine lactone ligation, actually forming an amide bond. We discuss in Chapter 1 studies of cytochrome *c* where CNBr cleavage generates two fragments that can reconstitute, catalysed by the proximity of the two reacting site termini

(Corradin and Harbury, 1974, Proudfoot et al., 1989). We wondered if the affinity reported in the literature for calbindin D_{9k} could be a result of the known side effect of CNBr cleavage at a Met-Ser site, the homoserine lactone formed being reactive under such conditions of proximity and generating an amide bond following attack by the Ser side chain. Therefore this value could represent a covalent rather than non-covalent interaction.

To investigate further the peptides were modified to contain a Ser-Ser interaction site. This was to serve as preliminary work before the more expensive homoserine was to be included, using pre-loaded Ser resin in its place. With these peptides we were able to identify a K_d value of 59.8 nM ± 12.8 nM (Figure 6.2). It seems that changing two residues at the interaction site has had little impact on the interaction of the peptides.

Following this a Gly-Gly interaction was studied, and with a K_d of 69.0 nM ± 9.8 nM (Figure 6.3) this confirms synthetic flexibility at the interaction site. It appears that as this site is situated in the loop region between EF1 and EF2 in the intact protein, mutating the residues at these reacting termini – at least to the extent that we have probed – has had little effect on the affinity of the peptides for one another. Subsequently a Gly-Gly interaction site was decided upon for further experiments, hypothesized to be the most likely site with which to successfully observe conformationally-assisted ligation.

The final BLI experiment repeated the Gly-Gly work but in an EDTA-containing buffer to establish the effect of the absence of calcium on the affinity. In this case the results showed non-specific binding of EF1 to the sensor, where the response traces of the control experiments mirrored those of the actual interaction (Figure 6.4).

CD spectroscopy was next used to study the conformation of these peptides prior to attempting their conformationally-assisted ligation. In the work described in the introduction to this chapter, fluorescence spectroscopy was also carried out using the Tyr of EF1 to monitor complex formation when wt-EF1 and wt-EF2 were mixed. An EF2 variant with and additional reporter by the substitution F66W was also used. They observed a linear change in fluorescence upon titration of the added peptide until the 1:1 stoichiometry was achieved, confirming that these peptides fold and form the complex upon interaction. They also used competition experiments to compare the competition between binding of EF2 variants and F66W-EF2 with wt-EF1. Three variants, including wt-EF2, displayed a similar affinity to F66W-EF2, which enabled us to use EF2 with Trp66. Though these results were less precise, they were in agreement with those of the SPR experiments.

The BLI results in section 6.2 demonstrate the importance of the role of calcium, with affinities in the low nM range in its presence and non-specific binding when a chelator was present. CD provided greater insight into the conformation of the peptides under both conditions (Figure 6.5). Each experiment gives four spectra, where first both peptides were analysed separately before mixing and re-investigating. In both sets of CD spectra, with and without calcium, minima at 208 nm and 222 nm are indicative of the α–helical content expected given the published structure. In the EDTA experiment the analysis of the peptides when mixed exactly matches that predicted based on their individual analyses. However in the calcium buffer the peptides when mixed appear to develop greater α–helical content than they possess when separate. Each peptide binds one calcium ion which coordinates structure, and it appears that the combination of this along with their interaction when mixed is conducive to further secondary structure

formation. This experiment demonstrates the effect of the presence of calcium in the reaction buffer, and is in agreement with the results of the BLI work.

Some difficulties were encountered during the synthesis of these peptides, particularly with EF2. It was found necessary to resuspend both peptides in denaturing buffer for purification by preparative-HPLC, as they formed a gel when HPLC buffers were used as standard. As determined by these CD results the peptides have sufficient secondary structure in the absence of calcium that their folding in HPLC buffers has reduced their solubility. As solubility is a problem future syntheses of this peptide could include aggregation-minimizing modifications as discussed in previous chapters, for example pseudoprolines, to improve the efficiency of the synthesis.

6.6.2 Conformationally-assisted peptide N-(Me)sulfonamide ligation

Though the conformationally-assisted ligation attempts with the other split proteins studied, RNase A and CI2, have not been particularly successful we were optimistic of greater success with the calbindin D_{9k} peptides. For conformationally-assisted ligation EF1 was synthesized as a peptide N-(Me)sulfonamide, in concurrence with the previous attempts using S-peptide N-(Me)sulfonamide and CI2(1-39) N-(Me)sulfonamide. The BLI and CD results demonstrate the effect of calcium on these peptides, so the non-denaturing buffer contained 2 mM $CaCl_2$ in the first ligation attempt. Here we see both the EF1 MPAA thioester and the ligation product after 1 h, and the reaction is complete by 8 h, with clean HPLC traces (Figure 6.6). To demonstrate the effect of the coordination of calcium ions the experiment was repeated with EDTA instead (Figure 6.7), and though the ligation could be seen to be completed

at similar times under both conditions it is clear that the 2 mM calcium present has increased the rate from the start and the reaction has reached its plateau sooner than in its absence. This is probably most easily visualised by comparing the 4 h traces for both experiments, where in the presence of 2 mM calcium both starting material peaks are reduced and the product peak is much more substantial (Figure 6.8).

In order to establish the true effect of secondary structure on this ligation, and to prove that it is in fact conformationally-assisted, the above experiment was repeated in the denaturing buffer. The presence of GdnHCl, even with calcium, should abolish any conformation of the peptides. After 24 h we see no change in the HPLC traces aside from the shift of the MPAA peak, both starting materials are still present and there is no evidence of product formation (Figure 6.9). There is a hint of a peak at an elution time corresponding to that of the ligation product but there is also a small peak in this area in the t = 0 time point, and perhaps is an impurity in the buffer. This confirms that the calbindin D_{9k} peptides used here are reliant on their secondary structure to drive ligation, and thus it is conformationally-assisted.

The identification of conditions where ligation could be "switched on and off" simply by the addition of calcium was of interest. As ligation was shown to occur to variable extents in the absence of calcium but in the presence of excess thiols, it was considered whether the addition of calcium to a thiol-free mixture could be enough to induce ligation. Though increasing the amount of MPAA increased the rate of ligation, the difference made by the addition of calcium could allow for ligation in the complete absence of any thiols. Two ligation attempts were set up in parallel, both using non-denaturing buffer but one with EDTA and one with $CaCl_2$. The EDTA-containing ligation experiment did not generate any product as anticipated, and neither did the calcium-containing experiment (Figure 6.10). This confirms that thiols are necessary, at

least when using EF1 N-(Me)sulfonamide. Future work could investigate the effect of synthesizing EF1 as alternative peptide thioesters, work that was not permitted in the time allowed for this project.

A ligation with MESNA (10 % w/v) as an alternative thiol was based upon published work generating MESNA thioesters with this amount (Adams et al., 2013). Thioester exchange was complete by 8 h, potentially at around 4 h in the experiment with calcium, though there is little ligation (Figure 6.11). Therefore MESNA can be ruled out as an appropriate thiol for this reaction. Future work could investigate the effect of other thiols on the rate of conformationally-assisted ligation, perhaps including benzyl mercaptan or thiophenol as previously used with the other split proteins studied herein.

Thus far the only ligation with both calcium and MPAA present has used 300 mM thiol. It was decided to reduce the MPAA to 60 mM. Ligation product is visible after 2 h, though only goes to completion overnight, by 24 h (Figure 6.12). This is significantly slower than the 8 h completion with 300 mM MPAA.

Having demonstrated conformationally-assisted ligation with these peptides at 1 mM the concentration of both starting materials was lowered to test the effect, reducing the concentration to just 100 μM. The problem of finding analytical-HPLC conditions to monitor such a reaction soon became apparent: injecting neat reaction mix overloaded the column with MPAA, but diluting it left the reactant peaks too small to identify against the baseline (Figure 6.13). One method by which this ligation could possibly be monitored is by running sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE), though this in itself contains potential problems. A higher percentage of acrylamide would be needed in order to increase cross-linking as the ligation product, being little over 8 kDa, is quite small by electrophoresis standards. However in doing

this the resolution of the bands will be much lower, adding to the difficulty in identifying the product. Another alternative could be to dialyse the sample to remove the MPAA, with mini dialysis tubes available for volumes as small as 10 µl with molecular weight cut-off (MWCO) as low as 2 kDa, and microdialysis plates for volumes as low as 10 µl and MWCO from 3.5 kDa. This would permit HPLC analysis of the time points, even when the ligation is conducted on a small scale as many in this chapter are.

The final experiment in this section explored the thiolysis step in the ligation, EF1 N-(Me)sulfonamide converting to EF1 MESNA thioester. Thioester exchange has been demonstrated in previous chapters with both S-peptide N-(Me)sulfonamide and CI2(1-39) N-(Me)sulfonamide, and having observed this exchange in the calbindin D_{9k} conformationally-assisted ligations a separate experiment was carried out to monitor it. Under the same conditions as in the ligation with MESNA EF1 N-(Me)sulfonamide was cleanly and successfully converted to EF1 MESNA thioester in just over 4 h (Figure 6.14). This is quicker than the same step in the prior MESNA ligation because there is no EF2 present for EF1 to interact with, making EF1 more readily accessible.

6.6.3 Proximity-assisted acyl transfer

Following the success of the conformationally-assisted ligations the focus then shifted to proximity-assisted acyl transfer. Chapter 1 contains examples of proximity-induced chemical reactions, notably the FLAG-tag and CA6D4 work (Nonaka et al., 2009, Nonaka et al., 2010). Additionally proximity-assisted acyl transfer has been performed with coiled-coils that mimic the aminoacyl transfer reactions of ribosomes

and non-ribosomal peptide synthetases, designed to increase rate by positioning the reacting components in close proximity (Leman et al., 2007). A proximity-assisted chemical reaction such as acyl transfer has not been attempted in previous chapters of with any of the other split proteins because their ligation has not been as successful as anticipated. We propose that the observed ligation in these systems was potentially sufficient for proximity-assisted acyl transfer. Here we examine the prospect for transfer of a tag from the *C*-terminus of EF1 to the *N*-terminus of EF2.

The first step was to modify EF1 N-(Me) sulfonamide for a *C*-terminal cysteamine using cysteamine hydrochloride (Figure 6.15). This allows loading of the tag onto EF1 before transfer to EF2. EF1 cysteamine and EF2 were mixed with a small peptide tag, LAPAG-MPAL thioester, at a 1:1:1 ratio. This tag was chosen for its small size, considered ideal for a first attempt. EF1 loaded with the tag is seen at 2 h, and the tag is the first component to be depleted. With EF1 cysteamine remaining more tag was added, and so at 5 h the total amount of tag relative to EF1 cysteamine was 10 eq. The first evidence of acyl transfer occurs in the 8 h time point where a new peak appears with an observed mass of 4085.8 Da by MALDI-TOF, corresponding to the calculated mass of 4092.47 for EF2 with the tag bound. A time point the next day at 26 h shows this peak has increased, though an additional peak with an observed mass of 4088.1 Da has also appeared (Figure 6.16). One explanation for this could be that the tag has bound to a site other than at the α–amino group. Future work to test this would be to cap the *N*-terminus of EF2 to prevent transfer of the tag to the expected site. If one peak still appears then this would indicate transfer of the tag to an unknown site.

6.6.4 Conformationally-assisted and Cu-free click chemistry

With successful conformationally-assisted at a Gly-Gly interaction site, section 6.5 sees the investigation into conformationally-assisted and Cu-free click chemistry, using the combination of affinity and structure to encourage a reaction that typically requires high temperature or the presence of a catalyst such as Cu to proceed. EF1 was synthesized with a *C*-terminal azide using the Fmoc-azido-NovaTag resin, and EF2 with an *N*-terminal alkyne by coupling propiolic acid at the end of the synthesis. Both are fully automated processes, achieved using microwave-assisted Fmoc-SPPS.

The first attempt did not appear successful and so to encourage ligation both the peptide concentration and the temperature were increased. With the peptides now at 10 mM and the temperature at 50 °C it appeared that the EF2 ΔM43/S44alkyne/F66W peak was smaller in the 2 h time point. In case the ligation product was eluting at a higher percentage of AcN the gradient was extended to 20 – 90 % B in 30 min. During the next HPLC injection, of the 4 h time point, we now see a clear third peak not visible before, corresponding to the ligation product (Figure 6.17). With this knowledge a shoulder on the EF2 ΔM43/S44alkyne/F66W peak is now apparent in the 2 h trace, corresponding to poor separation of the ligation product. Looking back to the first ligation attempt, solely analysed with this incorrect analytical-HPLC gradient, here too is evidence of a shoulder to the EF2 ΔM43/S44alkyne/F66W peak.

Though only monitored to 8 h the ligation is clearly progressing under these conditions. To explore whether the conditions for the first attempt had indeed been conducive to the reaction's progress the peptides were once again lowered to 1 mM. The temperature was however lowered to 37 °C rather than 30 °C, as the ligation was still not complete by 8 h at 50 °C. Ligation product is visible from 1 h, and a substantial amount in the 25 h time point (Figure 6.18). Note the 25 h HPLC trace is shifted with

respect to the previous time points as the HPLC injection loop was changed from 100 µl to 20 µl. Significantly for the possible application of this technique the buffer was not degassed prior to its addition to the peptides and no negative side effects were observed, with classic click chemistry performed under rigorously air-free conditions. From this point on the ligation buffer was never degassed for any click ligation. This makes the ligation conditions much easier, facilitating the ease of handling.

In a similar control experiment to that in section 6.3 a denaturing buffer was employed to determine if this ligation was not just Cu-free but conformationally-assisted. Under denaturing conditions no ligation product was observed by 24 h (Figure 6.19), and so we can conclude that the reaction does depend on the peptides having sufficient secondary structure.

These findings are encouraging and once established to occur, conditions were sought to improve the rate of the ligation. Returning to the calcium non-denaturing buffer the temperature was increased once more to 50 °C. A ligation under these conditions was only followed to 8 h in the earlier attempt, and ligation appeared complete by 24 h (Figure 6.20). These early attempts without Cu(I) as a catalyst are encouraging.

Having established the effect of temperature on the speed of the ligation, peptide concentration was lowered to observe its effect. With a peptide concentration of 0.5 mM ligation again appears complete in 24 h (Figure 6.21). As with previous ligation attempts EF2 ΔM43/S44alkyne/F66W is depleted before EF1 azide, and so the next ligation reduced the amount of EF1 azide to 0.25 mM. Once again the ligation is complete in 24 h (Figure 6.22). At this point the 8h analytical-HPLC traces from the last three experiments are compared: 1 mM peptides at 50 °C, 0.5 mM peptides at 50 °C and 0.25 mM EF1 azide with 0.5 mM EF2 ΔM43/S44alkyne/F66W-NH_2 at 50 °C. In each

experiment the ratio of all three peaks is comparable, confirming that the ligation is not concentration-dependent in this range (data not shown).

The work by Mock and colleagues discussed in Chapter 1 with a cucurbituril additive was conducted at 40 °C and provided an 11-fold enhancement to the rate of triazole formation (Mock et al., 1983, Mock et al., 1989). With half the amount of EF1 azide relative to EF2 ΔM43/S44alkyne/F66W-NH_2 the temperature was lowered to 45 °C. Under these conditions the ligation is now not complete until 48 h (Figure 6.23). To test reproducibility, and to confirm the concentrations of both starting materials, this ligation was repeated with concentrations determined by NanoDrop.

The final click experiment used NMR spectroscopy (carried out by Dr Geoff Kelly) to probe for which triazole isoform is present. As discussed in Chapter 1 Cu(I)-catalysed AAC solely results in the 1, 4-regioisomer (Tornoe et al., 2002, Rostovtsev et al., 2002), but as our reaction is Cu(I)-free we wanted to confirm which isomer had formed. The literature states that the protonated carbon of the two isomers occurs at different chemical shifts, with that of a 1,4-distributed 1,2,3-triazole at around 120 ppm and that of a 1,5-distributed 1,2,3-triazole at around 130 ppm (Creary et al., 2012). These signals occur in the aromatic region of a $^{1}H/^{13}C$-HSQC spectrum, and with the ~120 ppm region for our calbindin D_{9k} product containing just the one signal from the single Tyr and the five signals from the introduced Trp we should be able to identify a signal for the 1,4-regioisomer if present. Due to insufficient peptide concentration the signal-to-noise of the recorded $^{1}H/^{13}C$-HSQC spectrum is poor, though at least seven signals are observed in the ~120 ppm region (Figure 6.24). This extra signal is indicative of the 1,4-regioisomer having formed, though the experiment will have to be repeated with more sample to improve signal-to-noise. Additionally denaturing conditions could be used to unfold the click product, generating a sharper spectrum with

peaks from aromatic resides in random coil formation to improve signal assignment. This applies not only to the ~120 ppm region for identification of the 1,4-regioisomer but also to the ~130 ppm region to probe for a signal from the 1,5-regioisomer. Without accurate signal assignment in these regions we can at this point only conclude that the presence of the 1,4-regioisomer is indicated, future work establishing if the 1,5-regioisomer is also present.

This work has successfully demonstrated conformationally-assisted and Cu-free click chemistry using the calbindin D_{9k} peptides. We have established the effect of temperature on rate, and thus far not observed a decrease in rate when lowering the starting material concentrations. Future work could study this further, lowering the concentration to 100 μM as in the conformationally-assisted ligation using EF1 N-(Me)sulfonamide. Unlike that ligation, here we would not have to be concerned with thiols overloading the column, in theory allowing for HPLC analysis of lower and lower peptide concentrations.

Propiolic acid was chosen as the *N*-terminal alkyne component of EF2 based on work by Meldal and colleagues who used it for CuAAC on solid phase (Tornoe et al., 2002). In this instance the alkyne is a Michael acceptor that could be encouraging triazole formation. To investigate this future work could include alkyne variants such as the strained alkynes used in Cu-free click reactions mentioned in Chapter 1. A first attempt has been conducted using dibenzocyclooctyne-acid, though appropriate conditions for this reaction have yet to be determined as a result of time constraints. Future work could investigate alternative reaction azide and strained alkyne components. Although encouraging, the rate of the reaction is still too low to be an effective method for protein labelling *in vivo*. However, as demonstrated here it can

significantly assist a reaction that doesn't occur in its absence, suggesting that it could accelerate reactions that are currently too slow to be effective for protein labelling (Lang and Chin, 2014).

Future work applicable to all ligations discussed within this chapter could investigate interactions aside from just Gly-Gly. Gly-Met and Ser-Ser were observed to possess comparable affinities and demonstrate a level of flexibility to the amino acids at the reaction site, and so other ligation junctions could be studied. Additionally a homoserine lactone could be added to the *C*-terminus of EF1 to see what effect this will have, in case it can form an amide bond upon proximity to the *N*-terminal Ser of EF2.

6.7 Summary

Calbindin D_{9k} was chosen as the final split protein, with the synthesis of EF1 and EF2 variants to create a sterically-unhindered Gly-Gly interaction site. We measured a K_d of 69.0 nM between the peptides, significantly different to the 3 pM affinity which had attracted us to this system initially. The discrepancy was probably caused by relegation of the homoserine lactone. The presence of 2 mM calcium improved conformationally assisted peptide N-(Me)sulfonamide ligation by enhancing α–helical structure, as confirmed by CD. Finally these peptides were demonstrated to perform conformationally-assisted and Cu-free click chemistry between a peptide azide and alkyne, establishing this system as the most promising for our proposed labelling method.

7. Concluding remarks

The aim of the project was to explore the use of split proteins in a novel protein labelling method. We would express a small portion of a split protein at the *N*-terminus of the protein of interest, which would be selectively targeted by its complementing split protein fragment carrying the label. The inherent structure of the split protein fragments will be used to induce a conformationally-assisted ligation, where the two peptides become covalently bound and the target protein is labelled *via* this complex. Chapter 1 discusses the wide range of labelling methods available from the classic GFP to site-directed mutagenesis for covalent labelling to Cys residues. The more recent enzyme-, affinity- and proximity- based methods that have been developed show that bioorthogonal labelling is an expanding area of work of great interest interest (Lang and Chin, 2014).

In each split protein system the conformationally-assisted ligation of synthetically prepared peptide fragments was the first point of study, to confirm the suitability of their particular interaction site for use in the proposed labelling method. In place of synthesizing peptides with a *C*-terminal thioester we demonstrated the direct use of a peptide N-(Me)sulfonamide in NCL, discussed in Chapter 3. Following model ligation BPTI peptides demonstrated ligation comparable to that of NCL, without the need for a separate thioester exchange step. This performance led us to their use in the investigation of conformationally-assisted ligation in the split protein systems studied.

RNase A was chosen as the first split protein system for investigation, being a classic split protein system and the first demonstrated example of reconstitution. We began by attempting to demonstrate conformationally-assisted ligation between a synthesized S-peptide N-(Me)sulfonamide and commercially-available S-protein.

Though ligation occurred we encountered a problem in the extent of hydrolysis of S-peptide N-(Me)sulfonamide, which we attempted to rectify by experimenting with thioesters of different reactivity. Oxidation of the thiol also proved problematic due to the absence of reducing agent in the buffer, as TCEP-HCl could not be included in the ligation buffer because of the disulfide bonds in the S-protein required for conformationally-assisted ligation. In addition to this we conducted the thioester exchange step separately to obtain the pre-formed S-peptide thioester, though ligation was still limited. Without obtaining ideal conditions the direction of study turned to alternative S-peptide variants to partner S-protein. BLI was used to establish the affinity of the full-length S-peptide, the truncated 15-mer peptide, the high-binding LB2 variant, and a variant with a *C*-terminal Gly for S-protein. The lowest K_d observed was 450 nM, belonging to the LB2 interaction with LB2, compared to the bench mark of 730 nM for the full-length S-peptide.

Though the result for the full-length S-peptide is in agreement with the Kossiakoff and co-workers, the K_d for RNase S varies in the literature depending on technique and conditions used to establish it. For example Richards and Vithayathil observed a 5 nM affinity (Richards and Vithayathil, 1959), whereas Woodfin and Massey identified an affinity of 20 μM (Woodfin and Massey, 1968). Discussions with co-workers suggest that the 599 nM K_d with the full-length S-peptide established by Kossiakoff and colleagues appears reasonable based on data obtained, however their LB2 data are not consistent with a 5 nM (personal communication with Dr Steve Martin). We also wondered if the discrepancy between our BLI results and published affinities may be due to immobilization of the S-peptide to the sensor. Kossiakoff and colleagues used SPR which should give comparable data to BLI, but they immobilized S-protein. S-protein is much larger so perhaps its *N*-terminus was more readily

accessible to S-peptide than the *C*-terminus of S-peptide was to S-protein in our experiments.

The work described on the study of RNase A has not replicated that published. A number of variations were made but instead of spending more time on this system we decided to move on to CI2, another well-studied split protein system. The literature reports a K_d value of 40 nM (de Prat Gay et al., 1994), and the peptides have undergone conformationally-assisted ligation (Beligere and Dawson, 1999). Conformationally-assisted ligation is first attempted at a Gly-Met interaction site, with CI2(1-39)T39G N-(Me)sulfonamide and CI2(40-64). A Gly is substituted at the *C*-terminus of CI2(1-39) to minimise steric hindrance after published data demonstrated conformationally-assisted ligation with the native Thr replaced with Asp. No ligation between the peptides is observed at this Gly-Met site, but conditions for thioester exchange with benzyl mercaptan are determined. As with RNase A we pre-exchanged for a thioester peptide and repeated ligation, but were still unable to replicate the published results. Fluorescence spectroscopy confirmed the peptides were interacting, with a concentration-dependent decrease in fluorescence. The next ligations substitute the native Met at the *N*-terminus of CI2(40-64) for Cys, and ligations with CI2(1-39)T39G N-(Me)sulfonamide are complete in 24 h with MPAA and just 2 h with thiophenol. The same ligations under denaturing conditions did not proceed, confirming the conformationally-assisted nature of the reaction.

The introduction of a Cys at this site and the following ligation demonstrates the ability of these peptides for conformationally-assisted ligation, but not at a Cys-free site. We wondered if the *C*-terminal Asp in the published conformationally-assisted ligation may increase ligation rates by way of an anhydride, resulting from the *C*-terminal Asp. The succinic anhydride formed is more reactive than the initial thioester and can ligate

quicker when in proximity to the *N*-terminus of the partner peptide as a result of their conformation. We considered synthesizing CI2(1-39)T39D N-(Me)sulfonamide to try to replicate the published conformationally-assisted ligation but decided that even if we observed a comparable affinity to that published, this peptide would not be appropriate for the labelling method. The activity of the anhydride could potentially result in labelling of other *N*-termini besides that targeted, and we would lose the chemoselectivity provided by a conformationally-assisted reaction. A control to test for this would be to perform ligation in the presence of other peptides or proteins, but as later publications disputed the suitability of an Asp-Cys ligation site it was decided move on to a third split protein system.

Calbindin D_{9k} was chosen as its peptide fragments are reported to interact with a high affinity, and neither fragment is of too large a size for consideration in our proposed labelling method. A number of variants on these peptides were synthesized by Fmoc-SPPS, to investigate three sets of reacting site termini. Changing these amino acids was not found to significantly affect the affinity of the peptides for one another, and so with a K_d of 69.0 nM as established by BLI a Gly-Gly interaction site was used for further experimentation. These peptides performed conformationally-assisted ligation in 8 h at low millimolar concentrations. An initial attempt at proximity-assisted acyl transfer of a small peptide tag between these peptides gave promising results, and the conditions will need to be further studied to identify those best suited.

In addition to conformationally-assisted ligation with a peptide N-(Me)sulfonamide these peptides demonstrated conformationally-assisted and Cu-free click chemistry. EF1 was synthesized with a *C*-terminal azide, and EF2 with an *N*-terminal alkyne, and at relatively low temperatures and low millimolar concentrations

ligated to form the triazole-linked product. An initial attempt at establishing the triazole isoform was conducted using $^{1}H/^{13}C$-HSQC based on work by Funk and co-workers (Creary et al., 2012). The results of this first experiment were not reliable enough to draw conclusions though a signal in the ~120 ppm region indicates the 1,4-disubstituted-1,2,3-triazole, and so this will have to be repeated with product collected from a ligation of larger scale.

Ultimately we intend to continue working with calbindin D_{9k} to develop the best system for protein labelling by optimising the peptide components. The work described in Chapter 6 is highly promising for the development of a proximity-assisted labelling method, with impressive conformationally-assisted ligations. Future work will involve the investigation of alternative azide and alkyne components, reducing peptide size without compromising affinity to allow for incorporation of the smallest peptides possible into the proposed labelling method, and exploring non-canonical amino acids. There are a number of routes by which we can optimise this interaction, and improve upon these exciting initial results.

The experiments performed thus far have concentrated on establishing the ideal split protein system for these conformationally-assisted chemical reactions. If the work with RNase A and CI2 had not required such iterative condition changes we would have concluded with the first demonstration of *in vivo* protein labelling, or at least *in vitro* labelling. Calbindin D_{9k} proved to be the most exciting candidate, combining relatively small peptides with high affinity, and so we will begin labelling investigation with this split protein. In addition to its peptide lengths reconstitution of calbindin D_{9k} results in a calcium-transporter complex and not a functional ribonuclease as with RNase A, for example, and so is better suited for cellular labelling. In the proposed method EF1

would be the donor peptide carrying the label or tag and EF2 would be the acceptor peptide expressed on the target protein. If the method of ligation using a peptide N-(Me)sulfonamide were to be employed we would first need to pre-exchange to the thioester peptide, so as to not add thiols to the cellular environment. However the click azide and alkyne components have been shown to be physiologically compatible (Prescher and Bertozzi, 2005), and so this form of ligation could be incorporated into the labelling method. The first step would be to perform *in vitro* labelling of a target protein, by expressing EF2 as the acceptor peptide at the protein's *N*-terminus. The first target for consideration would be a small ubiquitin-like modifier (SUMO) protein, as these are small in size and labelling could be monitored by HPLC (Marblestone et al., 2006). A larger protein target will also be investigated, for though SUMO labelling should be easy to monitor we would need to ensure that the tag was not perturbing its activity due to its size. Maltose-binding protein (MBP) is a commonly used solubility enhancer (Fox and Waugh, 2003), and could be an ideal candidate for demonstration of labelling. With either target we could begin with biotin as the tag, using streptavidin pull-down for isolation and identification of potential labelled target protein.

Once we have demonstrated *in vitro* labelling the goal would be to label a cellular protein. We have considered beginning with a muscarinic G-protein coupled receptor (GPCR), as its extracellular *N*-terminus will provide easier labelling for initial investigation. Expression of the acceptor peptide at this exosite allows for both the agonist and antagonist binding sites to remain free, retaining full signalling activity of the receptor. This also allows for tests to confirm the successful expression of the modified receptor at the cell membrane by using the fluorescent antagonist Cy3B-telenzepine (Hern et al., 2010). Live cell single-molecule imaging could be used to estimate the level of expression of the modified muscarinic receptor, and it should be

possible to use plasmids encoding the muscarinic receptor fused to GFP by replacing the GFP sequence with the desired peptide sequence. CHO cells are one cell type for consideration because these do not express any muscarinic receptors on the plasma membrane, so any fluorescent molecules we can visualise during the imaging experiments will be the modified receptors. The initial fluorophore for testing will be Cy3B, a bright fluorophore that will make it easier to visualise the labelled receptors. Cy3B is brighter than GFP (Nenasheva et al., 2011) and alongside the specific affinity of the donor peptide for the small expressed acceptor peptide this method could prove to be a new and very powerful labelling technique.

References

ADAMS, A. L., COWPER, B., MORGAN, R. E., PREMDJEE, B., CADDICK, S. & MACMILLAN, D. 2013. Cysteine Promoted C-Terminal Hydrazinolysis of Native Peptides and Proteins. *Angewandte Chemie International Edition,* 52**,** 13062-13066.

AGARD, N. J., PRESCHER, J. A. & BERTOZZI, C. R. 2004. A Strain-Promoted [3 + 2] Azide-Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *Journal of the American Chemical Society,* 126**,** 15246-15047.

AGNEW, H. D., ROHDE, R. D., MILLWARD, S. W., NAG, A., YEO, W.-S., HEIN, J. E., PITRAM, S. M., TARIQ, A. A., BURNS, V. M., KROM, R. J., FOKIN, V. V., SHARPLESS, K. B. & HEATH, J. R. 2009. Iterative In Situ Click Chemistry Creates Antibody-like Protein-Capture Agents. *Angewandte Chemie International Edition,* 48**,** 4944-4948.

ANFINSEN, C. B., REDFIELD, R. R., CHOATE, W. L., PAGE, J. & CARROLL, W. R. 1954. Studies on the Gross Structure, Cross-Linkages and Terminal Sequences in Ribonuclease. *Journal of Biological Chemistry,* 207**,** 201-210.

ANTOS, J. M., CHEW, G. L., GUIMARAES, C. P., YODER, N. C., GROTENBREG, G. M., POPP, M. W. L. & PLOEGH, H. L. 2009. Site-Specific N- and C-Terminal Labeling of a Single Polypeptide Using Sortases of Different Specificity. *Journal of the American Chemical Society,* 131**,** 10800-+.

BACSA, B., DESAI, B., DIBÓ, G. & KAPPE, C. O. 2006. Rapid solid-phase peptide synthesis using thermal and controlled microwave irradiation. *Journal of Peptide Science,* 12**,** 633-638.

BASKIN, J. M., PRESCHER, J. A., LAUGHLIN, S. T., AGARD, N. J., CHANG, P. V., MILLER, I. A., LO, A., CODELLI, J. A. & BERTOZZI, C. R. 2007. Copper-free click chemistry for dynamic in vivo imaging. *Proceedings of the National Academy of Sciences,* 104**,** 16793-16797.

BELIGERE, G. S. & DAWSON, P. E. 1999. Conformationally assisted protein ligation using C-terminal thioester peptides. *Journal of the American Chemical Society,* 121**,** 6332-6333.

BELIGERE, G. S. & DAWSON, P. E. 2000. Design, Synthesis, and Characterization of 4-Ester CI2, a Model for Backbone Hydrogen Bonding in Protein α-Helices. *Journal of the American Chemical Society,* 122**,** 12079-12082.

BERGGÅRD, T., JULENIUS, K., OGARD, A., DRAKENBERG, T. & LINSE, S. 2001. Fragment complementation studies of protein stabilization by hydrophobic core residues. *Biochemistry,* 40**,** 1257-1264.

BLAKELEY, B. D., CHAPMAN, A. M. & MCNAUGHTON, B. R. 2012. Split-superpositive GFP reassembly is a fast, efficient, and robust method for detecting protein–protein interactions in vivo. *Molecular BioSystems,* 8**,** 2036.

BODANSZKY, M. & KWEI, J. Z. 1978. Side Reactions in Peptide-Synthesis .7. Sequence Dependence in Formation of Aminosuccinyl Derivatives from Beta-Benzyl-Aspartyl Peptides. *International Journal of Peptide and Protein Research,* 12**,** 69-74.

BURLINA, F., MORRIS, C., BEHRENDT, R., WHITE, P. & OFFER, J. 2012. Simplifying native chemical ligation with an N-acylsulfonamide linker. *Chemical Communications,* 48**,** 2579.

BURLINA, F., PAPAGEORGIOU, G., MORRIS, C., WHITE, P. D. & OFFER, J. 2014. In situ thioester formation for protein ligation using α-methylcysteine. *Chemical Science,* 5**,** 766-770.

CABANTOUS, S., NGUYEN, H. B., PEDELACQ, J.-D., KORAÏCHI, F., CHAUDHARY, A., GANGULY, K., LOCKARD, M. A., FAVRE, G., TERWILLIGER, T. C. & WALDO, G. S. 2013. A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association. *Scientific Reports,* 3.

CABANTOUS, S., TERWILLIGER, T. C. & WALDO, G. S. 2005. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology,* 23**,** 102-107.

CAREY, J., LINDMAN, S., BAUER, M. & LINSE, S. 2007. Protein reconstitution and three-dimensional domain swapping: Benefits and constraints of covalency. *Protein Science,* 16**,** 2317-2333.

CHAO, H. M., BAUTISTA, D. L., LITOWSKI, J., IRVIN, R. T. & HODGES, R. S. 1998. Use of a heterodimeric coiled-coil system for biosensor application and affinity purification. *Journal of Chromatography B,* 715**,** 307-329.

CHATTOPADHYAYA, R., MEADOR, W. E., MEANS, A. R. & QUIOCHO, F. A. 1992. Calmodulin Structure Refined at 1.7 Angstrom Resolution. *Journal of Molecular Biology,* 228**,** 1177-1192.

CORRADIN, G. & HARBURY, H. A. 1970. Cleavage of Cytochrome-C with Cyanogen Bromide. *Biochimica Et Biophysica Acta,* 221**,** 489-&.

CORRADIN, G. & HARBURY, H. A. 1971. Reconstitution of Horse Heart Cytochrome C - Interaction of Components Obtained Upon Cleavage of Peptide Bond Following Methionine Residue 65. *Proceedings of the National Academy of Sciences of the United States of America,* 68**,** 3036-&.

CORRADIN, G. & HARBURY, H. A. 1974. Reconstitution of Horse Heart Cytochrome-C - Reformation of Peptide-Bond Linking Residues 65 and 66. *Biochemical and Biophysical Research Communications,* 61**,** 1400-1406.

COTTON, G. J. & MUIR, T. W. 2000. Generation of a dual-labeled fluorescence biosensor for Crk-II phosphorylation using solid-phase expressed protein ligation. *Chemistry & Biology,* 7**,** 253-261.

CREARY, X., ANDERSON, A., BROPHY, C., CROWELL, F. & FUNK, Z. 2012. Method for Assigning Structure of 1,2,3-Triazoles. *The Journal of Organic Chemistry,* 77**,** 8756-8761.

CRICK, F. H. C. 1953. The Packing of α-Helices: Simple Coiled-Coils. *Acta. Cryst.,* 6.

DANG, B., KUBOTA, T., MANDAL, K., BEZANILLA, F. & KENT, S. B. H. 2013. Native Chemical Ligation at Asx-Cys, Glx-Cys: Chemical Synthesis and High-Resolution X-ray Structure of ShK Toxin by Racemic Protein Crystallography. *Journal of the American Chemical Society,* 135**,** 11911-11919.

DAWSON, P. E., CHURCHILL, M. J., GHADIRI, M. R. & KENT, S. B. H. 1997. Modulation of reactivity in native chemical ligation through the use of thiol additives. *Journal of the American Chemical Society,* 119**,** 4325-4329.

DAWSON, P. E., MUIR, T. W., CLARKLEWIS, I. & KENT, S. B. H. 1994. Synthesis of Proteins by Native Chemical Ligation. *Science,* 266**,** 776-779.

DE PRAT GAY, G. & FERSHT, A. R. 1994. Generation of a family of protein fragments for structure-folding studies. 1. Folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue. *Biochemistry,* 33**,** 7957-63.

DE PRAT GAY, G., RUIZ-SANZ, J., DAVIS, B. & FERSHT, A. R. 1994. The structure of the transition state for the association of two fragments of the barley chymotrypsin inhibitor 2 to generate native-like protein: Implications for mechinsms of protein folding. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 10943-10946.

DWYER, J. J., DWYER, M. A. & KOSSIAKOFF, A. A. 2001. High affinity RNase S-peptide variants obtained by phage display have a novel "hot-spot" of binding energy. *Biochemistry,* 40**,** 13491-13500.

ELMASRY, N. F. & FERSHT, A. R. 1994. Mutational Analysis of the N-Capping Box of the Alpha-Helix of Chymotrypsin Inhibitor-2. *Protein Engineering,* 7**,** 777-782.

FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature,* 340**,** 245-6.

FINN, B. E., KÖRDEL, J., THULIN, E., SELLERS, P. & FORSÉN, S. 1992. Dissection of Calbindin Dgk into two Ca2”-binding subdomains by a combination of mutagenesis and chemical cleavage. *FEBS Lett.,* 298**,** 211-214.

FOX, J. D. & WAUGH, D. S. 2003. Maltose-Binding Protein as a Solubility Enhancer. *Methods Mol Biol,* 205**,** 99-117.

FREEMAN, W. A., MOCK, W. L. & SHIH, N. Y. 1981. Cucurbituril. *Journal of the American Chemical Society,* 103**,** 7367-7368.

FUKUCHI, S. & NISHIKAWA, K. 2001. Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria. *Journal of Molecular Biology,* 309**,** 835-843.

GHOSH, I., HAMILTON, A. D. & REGAN, L. 2000. Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *J. Am. Chem. Soc.,* 122**,** 5658-5659.

GIRIAT, I. & MUIR, T. W. 2003. Protein semi-synthesis in living cells. *Journal of the American Chemical Society,* 125**,** 7180-7181.

GREENFIELD, N. & FASMAN, G. D. 1969. Computed Circular Dichroism Spectra for the Evaluation of Protein Conformation. *Biochemistry,* 8**,** 4108-4116.

GREENFIELD, N. J. 2006. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols,* 1**,** 2876-2890.

GROSS, E. & WITKOP, B. 1962. Nonenzymatic cleavage of peptide bonds: the methionine residues in bovine pancreatic ribonuclease. *J Biol Chem,* 237**,** 1856-60.

HAACK, T. & MUTTER, M. 1992. Serine Derived Oxazolidines as Secondary Structure Disrupting, Solubilizing Building-Blocks in Peptide-Synthesis. *Tetrahedron Letters,* 33**,** 1589-1592.

HACKENG, T. M., GRIFFIN, J. H. & DAWSON, P. E. 1999. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 10068-10073.

HANCOCK, W. S. & BATTERSBY, J. E. 1976. New Micro-Test for Detection of Incomplete Coupling Reactions in Solid-Phase Peptide-Synthesis Using 2,4,6-Trinitrobenzene-Sulphonic Acid. *Analytical Biochemistry,* 71**,** 260-264.

HANTGAN, R. R. & TANIUCHI, H. 1977. Formation of a Biologically-Active, Ordered Complex from 2 Overlapping Fragments of Cytochrome-C. *Journal of Biological Chemistry,* 252**,** 1367-1374.

HERN, J. A., BAIG, A. H., MASHANOV, G. I., BIRDSALL, B., CORRIE, J. E. T., LAZARENO, S., MOLLOY, J. E. & BIRDSALL, N. J. M. 2010. Imaging individual myosin molecules within living cells, methods in molecular biology. *PNAS,* 107**,** 2693-2698.

HIRATA, R., OHSUMI, Y., NAKANO, A., KAWASAKI, H., SUZUKI, K. & ANRAKU, Y. 1990. Molecular-Structure of a Gene, Vma1, Encoding the Catalytic Subunit of H+-Translocating Adenosine-Triphosphatase from Vacuolar Membranes of Saccharomyces-Cerevisiae. *Journal of Biological Chemistry,* 265**,** 6726-6733.

HIRS, C. H., MOORE, S. & STEIN, W. H. 1960. The sequence of the amino acid residues in performic acid-oxidized ribonuclease. *Journal of Biological Chemistry,* 235**,** 633-47.

HOLM, L., MOODY, P. & HOWARTH, M. 2009. Electrophilic Affibodies Forming Covalent Bonds to Protein Targets. *Journal of Biological Chemistry,* 284**,** 32906-32913.

HOLMGREN, A. 1972. Thioredoxin-C' - Reconstitution of an Active Form of Escherichia-Coli Thioredoxin from 2 Noncovalently Linked Cyanogen Bromide Peptide Fragments. *Febs Letters,* 24**,** 351-&.

HOLZWARTH, G. & DOTY, P. 1965. The Ultraviolet Circular Dichroism of Polypeptides. *Journal of the American Chemical Society,* 87**,** 218-228.

HOSSAN, A., ABU-MELHA, H., AL-OMAR, M. & AMR, A. 2012. Synthesis and Antimicrobial Activity of Some New Pyrimidinone and Oxazinone Derivatives Fused with Thiophene Rings Using 2-Chloro-6-ethoxy-4-acetylpyridine as Starting Material. *Molecules,* 17**,** 13642-13655.

HUISGEN, R. 1963. 1,3-Dipolar Cycloadditions Past and Future. *Angew. Chem. Int. Ed.,* 2**,** 565-632.

HYDE, C., JOHNSON, T., OWEN, D., QUIBELL, M. & SHEPPARD, R. C. 1994. Some Difficult Sequences Made Easy - a Study of Interchain Association in Solid-Phase Peptide-Synthesis. *International Journal of Peptide and Protein Research,* 43**,** 431-440.

INGENITO, R., BIANCHI, E., FATTORI, D. & PESSI, A. 1999. Solid phase synthesis of peptide C-terminal thioesters by Fmoc/t-Bu chemistry. *Journal of the American Chemical Society,* 121**,** 11369-11374.

ITZHAKI, L. S., OTZEN, D. E. & FERSHT, A. R. 1995. The Structure of the Transition State for Folding of Chymotrypsin Inhibitor 2 Analysed by Protein Engineering Methods: Evidence for a Nucleation-condensation Mechanism for Protein Folding. *J. Mol. Biol.,* 254**,** 260-288.

JACOB, E. & UNGER, R. 2007. A tale of two tails: why are terminal residues of proteins exposed? *Bioinformatics,* 23**,** e225-e230.

JEWETT, J. C. & BERTOZZI, C. R. 2010. Cu-free click cycloaddition reactions in chemical biology. *Chemical Society Reviews,* 39**,** 1272.

JOHNSON, E. C. B. & KENT, S. B. H. 2006. Insights into the mechanism and catalysis of the native chemical ligation reaction. *Journal of the American Chemical Society,* 128**,** 6640-6646.

JOHNSON, T., QUIBELL, M., OWEN, D. & SHEPPARD, R. C. 1993. A Reversible Protecting Group for the Amide Bond in Peptides - Use in the Synthesis of Difficult Sequences. *Journal of the Chemical Society-Chemical Communications***,** 369-372.

JOHNSSON, N. & VARSHAVSKY, A. 1994. Split Ubiquitin as a Sensor of Protein Interactions in-Vivo. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 10340-10344.

KAISER, E., COLESCOT.RL, BOSSINGE.CD & COOK, P. I. 1970. Color Test for Detection of Free Terminal Amino Groups in Solid-Phase Synthesis of Peptides. *Analytical Biochemistry,* 34**,** 595-&.

KAISER, R. & METZKA, L. 1999. Enhancement of cyanogen bromide cleavage yields for methionyl-serine and methionyl-threonine peptide bonds. *Analytical Biochemistry,* 266**,** 1-8.

KANG, J. & MACMILLAN, D. 2010. Peptide and protein thioester synthesis via N→S acyl transfer. *Organic & Biomolecular Chemistry,* 8**,** 1993.

KARAS, M., BACHMANN, D. & HILLENKAMP, F. 1985. Influence of the Wavelength in High-Irradiance Ultraviolet-Laser Desorption Mass-Spectrometry of Organic-Molecules. *Analytical Chemistry,* 57**,** 2935-2939.

KELEMEN, B. R., KLINK, T. A., BEHLKE, M. A., EUBANKS, S. R., LELAND, P. A. & RAINES, R. T. 1999. Hypersensitive substrate for ribonucleases. *Nucleic Acids Research,* 27**,** 3696-3701.

KENNER, G. W., MCDERMOTT, J. R. & SHEPPARD, R. C. 1971. The safety catch principle in solid phase peptide synthesis. *Journal of the Chemical Society D: Chemical Communications***,** 636.

KENT, S. B. H. 2009. Total chemical synthesis of proteins. *Chemical Society Reviews,* 38**,** 338.

KIICK, K. L., SAXON, E., TIRRELL, D. A. & BERTOZZI, C. R. 2001. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proceedings of the National Academy of Sciences,* 99**,** 19-24.

KOBAYASHI, N., HONDA, S., YOSHII, H., UEDAIRA, H. & MUNEKATA, E. 1995. Complement assembly of 2 fragments of the streptococcal protein-G B1 domain in aqueous-solution. *FEBS Lett.,* 366**,** 99-103.

KOJETIN, D. J., VENTERS, R. A., KORDYS, D. R., THOMPSON, R. J., KUMAR, R. & CAVANAGH, J. 2006. Structure, binding interface and hydrophobic transitions of Ca2+-loaded calbindin-D28K. *Nature Structural & Molecular Biology,* 13**,** 641-647.

KOLB, H. C., FINN, M. G. & SHARPLESS, K. B. 2001. Click chemistry: Diverse chemical function from a few good reactions. *Angewandte Chemie-International Edition,* 40**,** 2004-+.

KOSAKA, N., OGAWA, M., CHOYKE, P. L., KARASSINA, N., CORONA, C., MCDOUGALL, M., LYNCH, D. T., HOYT, C. C., LEVENSON, R. M., LOS, G. V. & KOBAYASHI, H. 2009. In Vivo Stable Tumor-Specific Painting in Various Colors Using Dehalogenase-Based Protein-Tag Fluorescent Ligands. *Bioconjugate Chemistry,* 20**,** 1367-1374.

KRETSINGER, R. H. & NOCKOLDS, C. E. 1973. Carp muscle calcium-binding protein. II. Structure determination and general description. *Journal of Biological Chemistry,* 248**,** 3313-26.

KUNITZ, M. 1946. A spectrophotometric method for the measurement of ribonuclease activity. *Journal of Biological Chemistry,* 164**,** 563-8.

LANG, K. & CHIN, J. W. 2014. Bioorthogonal Reactions for Labeling Proteins. *Acs Chemical Biology,* 9**,** 16-20.

LEE, D. L., GRANJA, J. R., MARTINEZ, J. A., SEVERIN, K. & GHADIRI, M. R. 1996. A self-replicating peptide. *Nature,* 382**,** 525-528.

LEMAN, L. J., WEINBERGER, D. A., HUANG, Z. Z., WILCOXEN, K. M. & GHADIRI, M. R. 2007. Functional and mechanistic analyses of biomimetic aminoacyl transfer reactions in de novo designed coiled coil peptides via rational active site engineering. *Journal of the American Chemical Society,* 129**,** 2959-2966.

LEWIS, W. G., GREEN, L. G., GRYNSZPAN, F., RADIC, Z., CARLIER, P. R., TAYLOR, P., FINN, M. G. & SHARPLESS, K. B. 2002. Click chemistry in situ: Acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. *Angewandte Chemie-International Edition,* 41**,** 1053-+.

LI, Z., SEOK, T. S. & JU, J. 2004. 1,3-Dipolar cycloaddition of azides with electron-deficient alkynes under mild condition in water. *Tetrahedron letters,* 45**,** 3143-3146.

LINSE, S., BRODIN, P., JOHANSSON, C., THULIN, E., GRUNDSTROM, T. & FORSEN, S. 1988. The Role of Protein Surface-Charges in Ion Binding. *Nature,* 335**,** 651-652.

LINSE, S., JOHANSSON, C., BRODIN, P., GRUNDSTROM, T., DRAKENBERG, T. & FORSEN, S. 1991. Electrostatic Contributions to the Binding of Ca2+ in Calbindin D9k. *Biochemistry,* 30**,** 154-162.

LOS, G. V., ENCELL, L. P., MCDOUGALL, M. G., HARTZELL, D. D., KARASSINA, N., ZIMPRICH, C., WOOD, M. G., LEARISH, R., OHANA, R. F., URH, M., SIMPSON, D., MENDEZ, J., ZIMMERMAN, K., OTTO, P., VIDUGIRIS, G., ZHU, J., DARZINS, A., KLAUBERT, D. H., BULLEIT, R. F. & WOOD, K. V. 2008. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol,* 3**,** 373-82.

LU, W., STAROVASNIK, M. A. & KENT, S. B. 1998. Total Chemical Synthesis of Bovine Pancreatic Trypsin Inhibitor by Native Chemical Ligation. *FEBS Lett.,* 429**,** 31-35.

LUDVIGSEN, S., SHEN, H., KJAER, M., MADSEN, J. C. & POULSEN, F. M. 1991. Refinement of the 3-Dimensional Solution Structure of Barley Serine Proteinase Inhibitor-2 and Comparison with the Structures in Crystals. *Journal of Molecular Biology,* 222**,** 621-635.

LUPAS, A. 1996. Coiled coils: new structures and new functions. *Trends in Biochemical Sciences,* 21**,** 375-82.

MAMIDYALA, S. K. & FINN, M. G. 2010. In situ click chemistry: probing the binding landscapes of biological molecules. *Chemical Society Reviews,* 39**,** 1252.

MARBLESTONE, J. G., EDAVETTAL, S. C., LIM, Y., LIM, P., ZUO, X. & BUTT, T. R. 2006. Comparison of SUMO fusion technology with traditional gene fusion systems: Enhanced expression and solubility with SUMO. *Protein Science,* 15**,** 182-189.

MENDE, F. & SEITZ, O. 2011. 9-Fluorenylmethoxycarbonyl-Based Solid-Phase Synthesis of Peptide α-Thioesters. *Angewandte Chemie International Edition,* 50**,** 1232-1240.

MEZZATO, S., SCHAFFRATH, M. & UNVERZAGT, C. 2005. An Orthogonal Double-Linker Resin Facilitates the Efficient Solid-Phase Synthesis of Complex-Type N-Glycopeptide Thioesters Suitable for Native Chemical Ligation. *Angewandte Chemie International Edition,* 44**,** 1650-1654.

MILLWARD, S. W., AGNEW, H. D., LAI, B., LEE, S. S., LIM, J., NAG, A., PITRAM, S., ROHDE, R. & HEATH, J. R. 2013. In situ click chemistry: from small molecule discovery to synthetic antibodies. *Integrative Biology,* 5**,** 87.

MOCK, W. L., IRRA, T. A., WEPSIEC, J. P. & ADHYA, M. 1989. Catalysis by Cucurbituril. The Significance of Bound-Substrate Destabilization for Induced Triazole Formation. *J. Org. Chem.,* 54**,** 5302-5308.

MOCK, W. L., IRRA, T. A., WEPSIEC, J. P. & MANIMARAN, T. L. 1983. Cycloaddition Induced by Cucurbituril - a Case of Pauling Principle Catalysis. *Journal of Organic Chemistry,* 48**,** 3619-3620.

MOOTZ, H. D. 2009. Split Inteins as Versatile Tools for Protein Semisynthesis. *ChemBioChem,* 10**,** 2579-2589.

MUIR, T. W. 2003. Semisynthesis Ofproteins Byexpressedproteinligation. *Annual Review of Biochemistry,* 72**,** 249-289.

MUIR, T. W., SONDHI, D. & COLE, P. A. 1998. Expressed protein ligation: A general method for protein engineering. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 6705-6710.

MÜLLER, J. & JOHNSSON, N. 2008. Split-Ubiquitin and the Split-Protein Sensors: Chessman for the Endgame. *ChemBioChem,* 9**,** 2029-2038.

MURALIDHARAN, V. & MUIR, T. W. 2006. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nature Methods,* 3**,** 429-438.

MURPHY, N. G., VARNEY, S. M., TALLON, J. M., THOMPSON, J. R. & BLANC, P. D. 2009. Fatal occupational exposure to trimethylsilyl-diazomethane. *Clinical Toxicology,* 47**,** 712.

NENASHEVA, T. A., MASHANOV, G. I., PECKHAM, M. & MOLLOY, J. E. 2011. Imaging Individual Myosin Molecules Within Living Cells. 778**,** 123-142.

NICOLAS, E., PEDROSO, E. & GIRALT, E. 1989. Formation of Aspartimide Peptides in Asp-Gly Sequences. *Tetrahedron Letters,* 30**,** 497-500.

NONAKA, H., FUJISHIMA, S.-H., UCHINOMIYA, S.-H., OJIDA, A. & HAMACHI, I. 2009. FLAG-tag selective covalent protein labeling via a binding-induced acyl-transfer reaction. *Bioorganic & Medicinal Chemistry Letters,* 19**,** 6696-6699.

NONAKA, H., FUJISHIMA, S., UCHINOMIYA, S., OJIDA, A. & HAMACHI, I. 2010. Selective Covalent Labeling of Tag-Fused GPCR Proteins on Live Cell Surface with a Synthetic Probe for Their Functional Analysis. *Journal of the American Chemical Society,* 132**,** 9301-9309.

OFFER, J., QUIBELL, M. & JOHNSON, T. 1996. On-resin solid-phase synthesis of asparagine N-linked glycopeptides: Use of N-(2-acetoxy-4-methoxybenzyl) (AcHmb) aspartyl amide-bond protection to prevent unwanted aspartimide formation. *Journal of the Chemical Society-Perkin Transactions 1***,** 175-182.

PAULING, L. & COREY, R. B. 1953. Compound helical configurations of polypeptide chains: structure of proteins of the alpha-keratin type. *Nature,* 171**,** 59-61.

PAULUS, H. 2000. Protein splicing and related forms of protein autoprocessing. *FEBS Lett.,* 69**,** 447-496.

POTTS, J. T., BERGER, A., COOKE, J. & ANFINSEN, C. B. 1962. A reinvestigation of the sequence of residues 11 to 18 in bovine pancreatic ribonuclease. *Journal of Biological Chemistry,* 237**,** 1851-5.

POTTS, J. T., JR., YOUNG, D. M. & ANFINSEN, C. B. 1963. Reconstitution of fully active RNase S by carboxypeptidase-degraded RNase S-peptide. *Journal of Biological Chemistry,* 238**,** 2593-4.

PRESCHER, J. A. & BERTOZZI, C. R. 2005. Chemistry in Living Systems. *Nature Chemical Biology,* 1**,** 13-21.

PROUDFOOT, A. E. I., ROSE, K. & WALLACE, C. J. A. 1989. Conformation-Directed Recombination of Enzyme-Activated Peptide-Fragments - a Simple and Efficient Means

to Protein Engineering - Its Use in the Creation of Cytochrome-C Analogs for Structure-Function Studies. *Journal of Biological Chemistry,* 264**,** 8764-8770.

QUIBELL, M., OWEN, D., PACKMAN, L. C. & JOHNSON, T. 1994. Suppression of Piperidine-Mediated Side Product Formation for Asp(Obu(T))-Containing Peptides by the Use of N-(2-Hydroxy-4-Methoxybenzyl) (Hmb) Backbone Amide Protection. *Journal of the Chemical Society-Chemical Communications***,** 2343-2344.

RAINES, R. T., MCCORMICK, M., VAN OOSBREE, T. R. & MIERENDORF, R. C. 2000. The S.Tag fusion system for protein purification. *Methods Enzymol,* 326**,** 362-76.

RICHARDS, F. M. 1958. On the Enzymic Activity of Subtilisin-Modified Ribonuclease. *Proc Natl Acad Sci U S A,* 44**,** 162-6.

RICHARDS, F. M. & VITHAYATHIL, P. J. 1959. The preparation of subtilisn-modified ribonuclease and the separation of the peptide and protein components. *Journal of Biological Chemistry,* 234**,** 1459-65.

ROSTOVTSEV, V. V., GREEN, L. G., FOKIN, V. V. & SHARPLESS, K. B. 2002. A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angewandte Chemie-International Edition,* 41**,** 2596-+.

RUDINGER, J. 1963. Organic chemistry in peptide synthesis. *Pure Appl. Chem.,* 7.

RYADNOV, M. G. & WOOLFSON, D. N. 2007. Self-assembled templates for polypeptide synthesis. *Journal of the American Chemical Society,* 129**,** 14074-14081.

SAGLE, L. B., ZIMMERMANN, J., DAWSON, P. E. & ROMESBERG, F. E. 2004. A high-resolution probe of protein folding. *Journal of the American Chemical Society,* 126**,** 3384-3385.

SAGLE, L. B., ZIMMERMANN, J., MATSUDA, S., DAWSON, P. E. & ROMESBERG, F. E. 2006. Redox-coupled dynamics and folding in cytochrome c. *Journal of the American Chemical Society,* 128**,** 7909-7915.

SANCHEZ-RUIZ, J. M., RAO, T., RUIZ-GÓMEZ, G., HILL, T. A., HOANG, H. N., FAIRLIE, D. P. & MASON, J. M. 2013. Truncated and Helix-Constrained Peptides with High Affinity and Specificity for the cFos Coiled-Coil of AP-1. *PLoS ONE,* 8**,** e59415.

SANTORO, J., GONZALEZ, C., BRUIX, M., NEIRA, J. L., NIETO, J. L., HERRANZ, J. & RICO, M. 1993. High-resolution three-dimensional structure of ribonuclease A in solution by nuclear magnetic resonance spectroscopy. *Journal of Molecular Biology,* 229**,** 722-34.

SAXON, E. 2000. Cell Surface Engineering by a Modified Staudinger Reaction. *Science,* 287**,** 2007-2010.

SCHROEDER, W. A., SHELTON, J. B. & SHELTON, J. R. 1969. An examination of conditions for the cleavage of polypeptide chains with cyanogen bromide: application to catalase. *Archives of Biochemistry and Biophysics,* 130**,** 551-6.

SCOTT, C. P., ABEL-SANTOS, E., WALL, M., WAHNON, D. C. & BENKOVIC, S. J. 1999. Production of cyclic peptides and proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 13638-13643.

SELA, M., WHITE, F. H., JR. & ANFINSEN, C. B. 1957. Reductive cleavage of disulfide bridges in ribonuclease. *Science,* 125**,** 691-2.

SEVERIN, K., LEE, D. H., KENNAN, A. J. & GHADIRI, M. R. 1997. A synthetic peptide ligase. *Nature,* 389**,** 706-9.

SEVERINOV, K. & MUIR, T. W. 1998. Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *Journal of Biological Chemistry,* 273**,** 16205-16209.

SHAH, N. H. & MUIR, T. W. 2014. Inteins: nature's gift to protein chemists. *Chemical Science,* 5**,** 446.

SHI, J. X. & MUIR, T. W. 2005. Development of a tandem protein trans-splicing system based on native and engineered split inteins. *Journal of the American Chemical Society,* 127**,** 6198-6206.

SHOEMAKER, K. R., FAIRMAN, R., KIM, P. S., YORK, E. J., STEWART, M. J. & BALDWIN, R. L. 1987. The C-peptide helix from ribonuclease A considered as an autonomous folding unit. *Cold Spring Harb Symp Quant Biol. ,* 52**,** 391-398.

SIMMONDS, R. G. 1996. Use of the Hmb backbone-protecting group in the synthesis of difficult sequences. *International Journal of Peptide and Protein Research,* 47**,** 36-41.

SMITH, V. F. & MATTHEWS, C. R. 2001. Testing the role of chain connectivity on the stability and structure of dihydrofolate reductase from E-coli: Fragment complementation and circular permutation reveal stable, alternatively folded forms. *Protein Science,* 10**,** 116-128.

SONG, H. & ISMAGILOV, R. F. 2003. Millisecond kinetics on a microfluidic chip using nanoliters of reagents. *Journal of the American Chemical Society,* 125**,** 14613-14619.

STAGLJAR, I., KOROSTENSKY, C., JOHNSSON, N. & TE HEESEN, S. 1998. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 5187-5192.

SUZUKI, T., OTA, Y., KASUYA, Y., MUTSUGA, M., KAWAMURA, Y., TSUMOTO, H., NAKAGAWA, H., FINN, M. G. & MIYATA, N. 2010. An Unexpected Example of Protein-Templated Click Chemistry. *Angewandte Chemie International Edition,* 49**,** 6817-6820.

SVENDSEN, I., JONASSEN, I., HEJGAARD, J. & BOISEN, S. 1980. Amino-Acid-Sequence Homology between a Serine Protease Inhibitor from Barley and Potato Inhibitor-1. *Carlsberg Research Communications,* 45**,** 389-395.

SVENSSON, L. A., THULIN, E. & FORSEN, S. 1992. Proline Cis-Trans Isomers in Calbindin D9k Observed by X-Ray Crystallography. *Journal of Molecular Biology,* 223**,** 601-606.

TANAKA, K., SHIROTSUKI, S., IWATA, T., KAGEYAMA, C., TAHARA, T., NOZAKI, S., SIWU, E. R. O., TAMURA, S., DOUKE, S., MURAKAMI, N., ONOE, H., WATANABE, Y. & FUKASE, K. 2012. Template-Assisted and Self-Activating Clicked Peptide as a Synthetic Mimic of the SH2 Domain. *Acs Chemical Biology,* 7**,** 637-645.

TANAKA, K., WAKI, H., IDO, Y., AKITA, S., YOSHIDA, Y. & YOSHIDA, T. 1988. Protein and Polymer Analyses up to mlz 100 000 by Laser Ionization Time-of-flight Mass Spectrometry. *Rapid communications in mass spectrometry,* 2**,** 151-153.

THOMAS, C. L. & MAULE, A. J. 2000. Limitations on the use of fused green fluorescent protein to investigate structure-function relationships for the cauliflower mosaic virus movement protein. *Journal of General Virology,* 81**,** 1851-1855.

THOMAS, F., BOYLE, A. L., BURTON, A. J. & WOOLFSON, D. N. 2013. A Set ofde NovoDesigned Parallel Heterodimeric Coiled Coils with Quantified Dissociation Constants in the Micromolar to Sub-nanomolar Regime. *Journal of the American Chemical Society,* 135**,** 5161-5166.

TORNOE, C. W., CHRISTENSEN, C. & MELDAL, M. 2002. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *Journal of Organic Chemistry,* 67**,** 3057-3064.

TSIEN, R. Y. 1998. The green fluorescent protein. *Annual Review of Biochemistry,* 67**,** 509-544.

TSUKIJI, S. & NAGAMUNE, T. 2009. Sortase-Mediated Ligation: A Gift from Gram-Positive Bacteria to Protein Engineering. *ChemBioChem,* 10**,** 787-798.

VENYAMINOV, S. U., BAIKALOV, I. A., SHEN, Z. M., WU, C.-S. H. & YANG, J. T. 1993. Circular dichroic analysis of denatured proteins: inclusion of denatures proteins in the reference set. *Analytical Biochemistry,* 214**,** 17-24.

VILLAIN, M., GAERTNER, H. & BOTTI, P. 2003. Native Chemical Ligation with Aspartic and Glutamic Acids asC-Terminal Residues: Scope and Limitations. *European Journal of Organic Chemistry,* 2003**,** 3267-3272.

WALLACE, C. J. A., CORRADIN, G., MARCHIORI, F. & BORIN, G. 1986. Cytochrome-C Chimerae from Natural and Synthetic Fragments - Significance of the Biological Properties. *Biopolymers,* 25**,** 2121-2132.

WALLACE, C. J. A. & PROUDFOOT, A. E. I. 1987. On the Relationship between Oxidation Reduction Potential and Biological-Activity in Cytochrome-C Analogs - Results from 4 Novel 2-Fragment Complexes. *Biochemical Journal,* 245**,** 773-779.

WANG, J., YU, Y. & XIA, J. 2014. Short Peptide Tag for Covalent Protein Labeling Based on Coiled Coils. *Bioconjugate Chemistry,* 25**,** 178-187.

WASSERMAN, R. H., CHANDLER, J. S., MEYER, S. A., SMITH, C. A., BRINDAK, M. E., FULLMER, C. S., PENNISTON, J. T. & KUMAR, R. 1992. Intestinal Calcium-Transport and Calcium Extrusion Processes at the Basolateral Membrane. *Journal of Nutrition,* 122**,** 662-671.

WAY, J. C. 2000. Covalent modification as a strategy to block protein-protein interactions with small-molecule drugs. *Current Opinion in Chemical Biology,* 4**,** 40-46.

WEHR, M. C., LAAGE, R., BOLZ, U., FISCHER, T. M., GRÜNEWALD, S., SCHEEK, S., BACH, A., NAVE, K.-A. & ROSSNER, M. J. 2006. Monitoring regulated protein-protein interactions using split TEV. *Nature Methods,* 3**,** 985-993.

WOHR, T. & MUTTER, M. 1995. Pseudo-Prolines in Peptide-Synthesis - Direct Insertion of Serine and Threonine Derived Oxazolidines in Dipeptides. *Tetrahedron Letters,* 36**,** 3847-3848.

WOMBACHER, R. & CORNISH, V. W. 2011. Chemical tags: Applications in live cell fluorescence imaging. *Journal of Biophotonics,* 4**,** 391-402.

WOODFIN, B. M. & MASSEY, V. 1968. Spectrophotometric determination of the dissociation constant of ribonuclease S'. *Journal of Biological Chemistry,* 243**,** 889-92.

WU, H., HU, Z. M. & LIU, X. Q. 1998. Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803. *Proceedings of the National Academy of Sciences of the United States of America,* 95**,** 9226-9231.

XUE, W.-F., CAREY, J. & LINSE, S. 2004. Multi-method global analysis of thermodynamics and kinetics in reconstitution of monellin. *Proteins: Structure, Function, and Bioinformatics,* 57**,** 586-595.

YAMAMURA, Y., HIRAKAWA, H., YAMAGUCHI, S. & NAGAMUNE, T. 2011. Enhancement of sortase A-mediated protein ligation by inducing a β-hairpin structure around the ligation site. *Chemical Communications,* 47**,** 4742.

YAO, S., GHOSH, I., ZUTSHI, R. & CHMIELEWSKI, J. 1997. A pH-modulated, self-replicating peptide. *Journal of the American Chemical Society,* 119**,** 10559-10560.

YI, L., SUN, H., ITZEN, A., TRIOLA, G., WALDMANN, H., GOODY, R. S. & WU, Y.-W. 2011. One-Pot Dual-Labeling of a Protein by Two Chemoselective Reactions. *Angewandte Chemie International Edition,* 50**,** 8287-8290.

ZAKERI, B., FIERER, J. O., CELIK, E., CHITTOCK, E. C., SCHWARZ-LINEK, U., MOY, V. T. & HOWARTH, M. 2012. Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** E690-E697.